A Model of Formative Assessment Practice in Secondary Science Classrooms using an

Audience Response System

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy in the Graduate School of

The Ohio State University

by

Melissa L. Shirley, M.S.

College of Education and Human Ecology

The Ohio State University

2009

Dissertation Committee:

Karen E. Irving, Advisor

Douglas T. Owens

Ross H. Nehm
Formative assessment involves the probing of students’ ideas to determine their level of understanding during the instructional sequence. Often conceptualized as a cycle, formative assessment consists of the teacher posing an instructional task to students, collecting data about student understanding, and engaging in follow-up strategies such as clarifying student understanding and adjusting instruction to meet learning needs. Despite having been shown to increase student achievement in a variety of classroom settings, formative assessment remains a relative weak area of teacher practice. Methods that enhance formative assessment strategies may therefore have a positive effect on student achievement.

Audience response systems comprise a broad category of technologies that support richer classroom interaction and have the potential to facilitate formative assessment. Results from a large national research study, Classroom Connectivity in Promoting Mathematics and Science Achievement (CCMS), show that students in algebra classrooms where the teacher has implemented a type of audience response system experience significantly higher achievement gains compared to a control group. This suggests a role for audience response systems in promoting rich formative assessment.
The importance of incorporating formative assessment strategies into regular classroom practice is widely recognized. However, it remains challenging to identify whether rich formative assessment is occurring during a particular class session. This dissertation uses teacher interviews and classroom observations to develop a fine-grained model of formative assessment in secondary science classrooms employing a type of audience response system. This model can be used by researchers and practitioners to characterize components of formative assessment practice in classrooms.

A major component of formative assessment practice is the collection and aggregation of evidence of student learning. This dissertation proposes the use of the assessment episode to characterize extended cycles of teacher-student interactions. Further, the model presented here provides a new methodology to describe the teacher’s use of questioning and subsequent classroom discourse to uncover student learning. Additional components of this model of formative assessment focus on the recognition of student learning by the teacher and the resultant changes in instructional practice to enhance student understanding.
“Education is learning what you didn’t even know you didn’t know”

Daniel J. Boorstin

I dedicate this work to all of the teachers and students who have influenced my thinking in formal classroom settings as well as informal interactions. I especially acknowledge teachers and advisors from the Merrimack School District (New Hampshire) who introduced me to the world, faculty from many universities who guided the development of my thinking, and colleagues and students of the Olentangy Local School District (Ohio) who taught me the most valuable lessons about learning.
Acknowledgments

This work would not have been possible without the assistance and support of many individuals, more than can be listed here. I hold you all in my thoughts.

I gratefully acknowledge the Principal Investigators of the Classroom Connectivity in Promoting Mathematics and Science Achievement (CCMS) research project: Doug Owens, Karen Irving, Stephen Pape, Louis Abrahamson, and Frank Demana. They have made my graduate experience and entry into the world of educational research much richer. I am grateful for their wisdom and experience, for the opportunity to participate in a large-scale research project, for allowing me to have access to data to begin my own studies, and for support for conference presentations. Many other members of this research team have been essential to my development as an academic and a researcher; I particularly thank Vehbi Sanalan for his tutelage, patience, collegiality, and friendship.

My work has also been guided by a number of faculty and staff members at The Ohio State University, especially members of my candidacy and dissertation committees: Karen Irving, Doug Owens, Ross Nehm, and Steve Rissing. Patti Brosnan and Gail Hoskins have played critical roles in my development as an educator of pre-service teachers. These people have given me teaching opportunities, independent study
experiences, job reference letters, and sage advice at various points in my studies. I am grateful to them all for their valuable assistance.

Eric Gordon, former principal at Olentangy Liberty High School, and Beth Lucek, my co-advisor of the Olentangy In The Know team, were instrumental in my decision to pursue my doctorate. I thank them for their friendship and support throughout our time together. Go Patriots!

My family has been so supportive throughout this endeavor. “Thank you” is inadequate to describe my appreciation for encouragement, study breaks, meals, yard work, computer services, help with moving, and most of all, their love.

Finally, it is difficult to express the extent of my gratitude to Karen Irving. From the beginning of my studies, she has guided my path as my faculty advisor and soon as my colleague. With words of encouragement and material support, she has made it possible for me to present our work at various conferences and to learn about the world of academic research.

The research reported here was supported by the Institute of Education Sciences, U.S. Department of Education, through Grant R305K050045 to The Ohio State University. The opinions expressed are those of the author and do not represent views of the U.S. Department of Education.
Vita

January 12, 1971 ........................................... Born in Binghamton, New York

1993 ......................................................... B.S., Carnegie Mellon

1997 ......................................................... M.S., Case Western Reserve University

2005 ......................................................... M.S., University of Dayton

1997-2006 .................................................. Science teacher, Olentangy Local Schools, Lewis Center, OH

2006-2007 .................................................. University Fellow, Ohio State University, Columbus OH

2007-2009 .................................................. Graduate Research Assistant, Ohio State University, Columbus OH

PUBLICATIONS


FIELDS OF STUDY

Major Field: Education
# Table of Contents

Abstract ......................................................................................................................... ii
Acknowledgments ......................................................................................................... v
Vita ................................................................................................................................. vii
Table of Contents ......................................................................................................... viii
List of Figures ............................................................................................................... xi
List of Tables ................................................................................................................ xii

## Chapter 1 : Introduction ......................................................................................... 1
- Statement of the Problem ......................................................................................... 3
  - Background .............................................................................................................. 3
  - Guiding Question ..................................................................................................... 4
  - Significance of the Study ........................................................................................ 5
- Structure of the Study ............................................................................................... 6
  - Definitions of Terms ............................................................................................... 6
  - Outline of the Dissertation .................................................................................... 9

## Chapter 2 : Theoretical Framework And Literature Review ............................... 11
- Formative Assessment .............................................................................................. 12
  - Definitions of Formative Assessment ...................................................................... 12
  - Models of Formative Assessment ............................................................................. 13
  - Components of Formative Assessment .................................................................. 25
- Interactive Classroom-Based Formative Assessment ............................................. 39
  - The Interactive Classroom ..................................................................................... 39
  - Audience Response Systems .................................................................................. 42
  - Interactive Networked Classroom Assessment Systems ....................................... 47
- Summary .................................................................................................................. 49

## Chapter 3 : Research Methodology - Intervention ............................................... 53
- Classroom Connectivity in Mathematics and Science Project ............................. 53
  - The Intervention .................................................................................................... 53
  - Quantitative Measures ............................................................................................ 59
  - Qualitative Measures .............................................................................................. 62
- Summary .................................................................................................................. 64
# Table of Contents

**Chapter 4 : Research Methodology – Data Analysis** ................................................................. 65

**Data Collection** .................................................................................................................. 66

Participants .............................................................................................................................. 66

**Data Analysis** .................................................................................................................... 69

Analysis of Interview Data ..................................................................................................... 69

Analysis of Student Observation Surveys ............................................................................... 72

Analysis of Classroom Observation Transcripts ..................................................................... 72

**Other Considerations** ........................................................................................................ 77

Trustworthiness ...................................................................................................................... 77

The Researcher as Instrument ................................................................................................. 80

**Summary** ............................................................................................................................ 81

**Chapter 5 : Results** ............................................................................................................ 82

**Defining Rich Formative Assessment Practice** .................................................................... 84

Assertion One: Rich Instructional Tasks are Implemented ....................................................... 84

Assertion Two: Questioning Is Deep and Probing ................................................................. 101

Assertion Three: Classroom Interactions Elicit Higher-Order Thinking .............................. 109

Assertion Four: Teachers Know What Their Students Understand .................................... 130

Assertion Five: Instructional Decisions are Based on Learning Evidence .......................... 137

**Summary** ............................................................................................................................ 147

**Chapter 6 : Discussion** ....................................................................................................... 150

**Summary of Results** ......................................................................................................... 151

Comparison of Model with Existing Models ................................................................. 151

Contributions of CCT to Rich Formative Assessment .......................................................... 158

**Significance and Implications for Practice** ......................................................................... 160

Significance of This Study ...................................................................................................... 160

Implications for Practice ........................................................................................................ 161

**Limitations of the Current Study** ........................................................................................ 166

Limited Numbers of Observations ....................................................................................... 166

No Student Collaborative Groups Recorded ......................................................................... 167

Representation of Typical Teacher Practice ....................................................................... 168

Limitations of the Data Set .................................................................................................... 169

Demonstration of Student Achievement .............................................................................. 170

Identifying Changes in Instruction ....................................................................................... 171

Limitations of Modeling ........................................................................................................ 172

**Future Work** ....................................................................................................................... 173

Changes in Formative Assessment Practice ......................................................................... 173

Enhancing Formative Assessment Practice ......................................................................... 176
List of Figures

Figure 2.1.  A conceptual model of formative assessment .............................................. 38
Figure 3.1.  Diagram of a classroom equipped with the TI-Navigator™ system .......... 56
Figure 5.1.  Use of instructional tasks .............................................................................. 101
Figure 5.2.  Degree of assessment episodes ....................................................................... 109
Figure 5.3.  Quality of assessment episodes in rich formative assessment .................... 130
Figure 5.4.  Awareness of student learning contributes to rich formative assessment ... 137
Figure 5.5.  Instructional decision-making ......................................................................... 147
Figure 5.6.  Model of rich formative assessment practice in whole-class discourse ..... 149
Figure 6.1.  Model of formative assessment .................................................................... 152
List of Tables

Table 3.1. Timeline of Intervention and Data Collection ........................................... 58
Table 4.1 Description of Selected Teachers .............................................................. 67
Table 4.2: Technology and Content Background of Participants ............................... 68
Table 4.3: Demographic Characteristics of School Settings ...................................... 68
Table 5.1 Science Topics Addressed in Instructional Tasks, by Participant .............. 86
Table 5.2. Purposes of Instructional Tasks Implemented Through CCT .................. 86
Table 5.3. Number and Duration of Assessment Episodes ........................................ 103
Table 5.4. Iterations of Assessment Episode by Teacher ........................................... 105
Table 5.5. Categorization of Initial Question Prompts by Participant ....................... 112
Table 5.6. Categorization of Subsequent Questions ................................................. 117
Table 5.7. Categorization of Responses in Assessment Episodes ............................. 123
Table 5.8. Student Survey Responses Regarding Teacher Knowledge of Learning .... 131
Table 6.1. Comparison of Formative Assessment Models ......................................... 157
Scientific literacy is an essential component of quality education for all students. A solid understanding of science is crucial for individuals to serve as responsible members of our global society. The American Association for Advancement in Science (Rutherford & Ahlgren, 1991) and the National Research Council (1996) suggest that scientific literacy is essential for individuals to understand the natural world as well as emerging technologies that allow nations to compete in the global arena. Furthermore, the study of science encourages a respect for the interactions within a given environment, promoting the notion that all life is somehow connected and interdependent. Environmental concerns benefit from this greater understanding as well as socio-political concerns: individuals may come to realize that their actions have repercussions in areas that may not seem immediately obvious (Rutherford & Ahlgren, 1991).

The critical thinking and problem solving skills that students develop in the study of science serve them regardless of their chosen careers. These skills are necessary to interpret the many competing claims of commercial and political interests inherent in society. Moreover, a strong foundation in scientific understanding for all students may increase the number of individuals who ultimately enter fields of study comprising the science, technology, engineering, and mathematics (STEM) disciplines. The resultant increase in scientific and technological advances in this country can then lead to positive
effects on the economy (Committee on Science, 2007; National Research Council, 1996; Rutherford & Ahlgren, 1991). Clearly, a strong understanding of science concepts and process is essential for full participation in today’s society and for the overall benefit of the United States and other nations.

Are students in the United States scientifically literate? Results from the National Assessment of Educational Progress (Grigg, Lauko, & Brockway, 2006; O'Sullivan, Lauko, Grigg, Qian, & Zhang, 2002) indicate that a majority of students in the United States are not at a proficient level of achievement in science. Findings from three years of administration of this assessment (1996, 2000, and 2005) demonstrate that approximately 30% of fourth grade and eight grade students and fewer than 20% of twelfth grade students meet this standard of scientific literacy (Grigg et al., 2006; O'Sullivan et al., 2002). Despite major efforts targeted at improving science education, average students have maintained this low degree of science literacy.

Are students in the United States, while performing at lower levels than science educators might like, outperforming their peers internationally? Two different international studies would suggest that this is not the case. According to the Programme for International Student Assessment, the United States lags behind other industrialized nations. In 2006, the international average score of 15-year-old students across 57 nations was 500 points. The average score for students from the United States was 489 points. Moreover, while on average, 19.2% of students attained the lowest two levels of performance, 24.2% of United States students performed no higher than the second level (Programme for International Student Assessment, 2007). In the 2003 Trends in International Mathematics and Science Study (Martin, Mullis, Gonzalez, & Chrostowski,
2004), eighth grade students in the United States performed at a level above the international average score (United States: 533, international average: 474). While above the international average performance, the U.S. is outperformed by nations that have a lower index of human development, such as Singapore, Estonia, and Hungary among others (Martin et al., 2004). Although the Trends in International Mathematics and Science Study result are more promising, they still support the notion that students in the United States do not have science achievement levels commensurate with the expectations or resources of this nation. Methods of teaching and learning science more effectively must continue to be studied and implemented.

Statement of the Problem

Background

Formative assessment has emerged as a promising instructional practice for increasing achievement and understanding in many content areas, including science. By making students’ thinking visible and allowing the teacher to gather a variety of information about student thinking, formative assessment strategies promote increased understanding and performance (Black & Wiliam, 1998a; Popham, 2008). Many different methods for conducting formative assessment exist, but one that holds great promise is interactive technology-facilitated assessment.

Researchers in a variety of fields have investigated the role of interactive classroom technology in promoting student achievement. The dissertation presented here includes data drawn from a larger research study investigating the role of interactive classroom-based formative assessment technology in promoting mathematics and science
achievement. Initial findings suggest that students enrolled in Algebra I classes equipped with such technology perform at higher levels on a mathematics achievement test than their counterparts without this technology (Pape et al., 2008). Additional work from this research study has attempted to understand aspects of teacher practice, including quantitative measures of implementation of interactive classroom technology (Irving et al., 2008) and qualitative descriptions of teachers’ experiences with such technology (Irving, Sanalan, & Shirley, 2009). Although this work, among many other studies (e.g. Crossgrove & Curran, 2008; Medina et al., 2008; Stroup, Carmona, & Davis, 2005), provides evidence for the benefit of interactive formative assessment technology in science classrooms, how these emerging technologies can best be used to promote student learning and how they fit with the variety of teachers’ classroom practices is not yet well-understood. The problem to be addressed in this dissertation is to characterize in detail aspects of formative assessment practice in secondary science classrooms where the teacher has implemented connected classroom technology.

Guiding Question

This study is designed to provide a detailed analysis of formative assessment as practiced by teachers in middle and high school science classrooms equipped with connected classroom technology. Specifically, it will address the following guiding question: How can rich formative assessment practice be described in detail? In answering this guiding question through the development of a new model for rich formative assessment practice, this dissertation will also address the following focus questions:
1. What instructional tasks provide rich assessment information?

2. How can classroom-based assessment episodes be analyzed in detail while retaining the context of the related science instruction?

3. What characteristics of assessment episodes support rich formative assessment practice?
   a. What types of questions are used at the beginning of assessment episodes?
   b. What types of student responses are elicited?
   c. What strategies are used in subsequent questions to probe student understanding?

4. How do teachers gather data about student learning?

5. How do teachers use formative assessment data to make instructional decisions?

**Significance of the Study**

The study presented in this dissertation is important for developing a deeper understanding of how the practice of formative assessment can be affected by the use of interactive technology. Technology is becoming increasingly prevalent in classrooms for all ages, and when used appropriately, can help teachers collect data used to make informed decisions regarding instructional practices. When equipped with more accurate and timely information, teachers may adapt instruction in ways that promote student learning. The particular form of connected classroom technology being examined in this dissertation can facilitate the collection, aggregation, and display of data concerning student understanding. This dissertation will examine the ways in which connected
classroom technology can affect the process of formative assessment and provide a model of formative assessment practice closely aligned with teacher practice.

This study goes beyond previous studies in several ways. First, many of the existing studies of interactive classroom technologies are conducted in large university lecture halls, rather than the secondary science classrooms used in the current dissertation. Second, unlike the “clicker” technology used in other studies, the technology that serves as part of the intervention in the current study allows for a wide range of instructional tasks to be administered. Finally, the participants in this study were not given prescribed lessons to complete; therefore, this study may identify ways in which different teachers adapted the technology to suit their particular situations. This can provide information regarding design of curricula and future technologies.

Structure of the Study

Definitions of Terms

Many of the terms used to describe science classrooms and student-teacher interactions are commonplace among science education researchers. However, this work will use certain terms related to assessment and classroom interactions in very specific ways. Further, this work will describe teacher use of a technology that will be unfamiliar to some readers. For these reasons, the following terms are defined as they are used in this study.

Secondary science. This term will be used collectively to describe science classes held for grades 7-12. This may include middle school general science as well as high school physical science, biology, chemistry, or physics courses.
Formative assessment. Formative assessment will be used broadly to describe the process of a teacher gathering information about student learning and acting on that information. Chapter 2 of this dissertation will describe in more detail the various components of formative assessment.

Task or use. Task and use are terms used in this dissertation to describe the prompt given to the students in order to assess their learning. A task or use is typically an activity initiated by the teacher, such as to have students complete a “bellringer” assignment, solve a physics problem, or collect and submit data from an experiment. Tasks are the ways in which a teacher chooses to implement formative assessment; in this study, many but not all tasks will be mediated through interactive classroom technology.

Follow-up. As part of the formative assessment cycle, teachers will follow-up on student responses in some way. Follow-ups might include the teacher assigning an additional problem, reviewing or reteaching a concept, issuing feedback to a student or to the entire class, or wrapping up that portion of a discussion and moving on.

Assessment episode. An assessment episode is the extended dialogue that occurs between the teacher and the students regarding a specific task or prompt. Follow-up questions from the teacher that probe student understanding or solicit additional responses are included as part of the same assessment episode as the initial task. An assessment episode will typically be anywhere from a brief question-and-answer exchange to an extended line of questioning lasting between five and ten minutes. If a teacher poses a supplemental sample problem for students or addresses a tangential issue, a new assessment episode would begin. A major emphasis of this dissertation is to characterize assessment episodes.
Technology-facilitated formative assessment. This term refers to any episode of the teacher learning more about student understanding through the use of technological devices. Technology-facilitated formative assessment can be mediated through several types of device, such as audience response systems including “clickers”; in this study, this process is implemented through the TI-Navigator™ system of CCT.

Audience response system (ARS). This is the general term for any device that allows responses from an audience (participants at a talk, students in a classroom) to be electronically collected and tabulated for the presenter or teacher. The most common form of an audience response system addressed in this dissertation is the “clicker.”

Clickers. These are small hand-held devices that wirelessly allow a student or audience member to submit a choice. Generally these are limited to ten or fewer choices and have numerical or multiple-choice functions only.

Connected classroom technology (CCT). This term will be used to refer specifically to a networked system of handheld devices such as the TI-Navigator™ system that allows the sharing of information between teacher and students. The distinction from an ARS is in the interactive nature of the system rather than a one-way submission of responses to question prompts.

Texas Instruments (TI)-Navigator™. This is a specific form of connected classroom technology that utilizes graphing calculators (TI 73, 83 Plus, 84 Plus families). The hardware and software components that comprise this system are described below.

Hardware components: The TI-Navigator™ uses wired and wireless components. Briefly, each student has his or her own graphing calculator plugged into a hub along with three other student calculators. The hub communicates wirelessly with an access
point connected to the teacher’s computer. The teacher’s computer is often connected to a digital projector so that information can be displayed to the entire class.

*Sofware components:* The TI-Navigator™ system includes five main functions. To distinguish one of these “components” from the actual pieces that make up the system, they will be referred to in this dissertation as “software components.” These software components include the ability to capture the text or images on each individual calculator screen, a variety of questioning applications, and software for the collection, aggregation, and display of data.

*Outline of the Dissertation*

In Chapter 2 of this dissertation, the literature regarding formative assessment, questioning, and audience response systems is briefly reviewed. Research-based models of formative assessment are presented first, followed by a synthesis of the various models and how they frame the dissertation study. The next major section of the literature review includes descriptions of quality question types and some of the more prominent question typologies. Finally, the literature review concludes with an overview of audience response technologies and classroom-based studies that examine their influence on teaching practice or student achievement.

The data for this dissertation study are taken from a larger research project, Classroom Connectivity in Promoting Mathematics and Science Achievement (CCMS). In Chapter 3, the design of the CCMS study and relevant data collection methods are described. Background information on the particular CCT system implemented in classrooms as part of the CCMS study is provided, followed by a discussion of the quantitative surveys and qualitative interviews and observations used for data collection.
The methodology for this particular dissertation study is then described in Chapter 4. Summary descriptive statistics about the demographics of the selected participants are presented. The various quantitative and qualitative measures that were administered to the participants are then described.

Chapter 5 presents the results of this dissertation study. Aspects of formative assessment practice are broken into components of assigning high-quality instructional tasks, gathering and interpreting data through questioning, teachers knowing about student learning, and teachers using gathered data about student learning to inform teaching practices. Each of these areas will be presented along with supporting evidence for the varieties of classroom practice that support different levels of formative assessment practice. Chapter 5 culminates in a model for rich formative assessment practice informed by the data gathered and analyzed in this study.

In Chapter 6, an overview of the dissertation study and a summary of the findings are presented. The model described in Chapter 5 will be discussed in light of others’ work regarding formative assessment. Implications for practice relevant to various audiences are introduced. Emergent issues related to methodological strengths and weaknesses are also addressed in this section. Chapter 6 will conclude with a description of the areas in which future research related to the findings of this dissertation are perceived to be especially fruitful and necessary.
Chapter 2: Theoretical Framework And Literature Review

The role of formative assessment (FA) in promoting student achievement has been clearly established by a number of researchers in various disciplines. In this chapter, findings and models of formative assessment from major researchers studying science classrooms will be described. Formative assessment consists of four primary facets: the administration of some type of instructional task to learners, teacher questioning to probe student understanding, an awareness of the degree of student understanding, and follow-up by the teacher in the form of feedback and/or changes in instructional strategy.

Further, the role of interactive classroom technologies in promoting student engagement and learning will be reviewed. Interactive classroom technologies include wireless “clicker” systems used for polling students in large lecture halls as well as networked handheld calculators. Connected classrooms using networked calculators have many of the facets of clicker systems, but add functionalities that allow for more open tasks and components that are better suited to mathematics and science classrooms. The concepts addressed in this review inform the design of the dissertation study described in the remainder of this document.
Formative Assessment

Definitions of Formative Assessment

Assessment has risen to the forefront of teacher practice, administrative policy, and educational research. As federal funds mandate achievement testing, many districts in turn implement regular batteries of achievement tests, which necessitate the teachers’ attention to ongoing student learning.

Four major types of assessment have been described (Wiliam, 2008). The first is monitoring assessment, where the teacher pays attention to trends to notice that a student is not doing well at a given task or concept. This form of assessment merely identifies that a problem may exist but it does not identify what the problem might be, nor does it suggest potential solutions. Second, diagnostic assessment is used to identify what the learning problem is, but also stops short of identifying ways that the learner could improve his or her work. Third, formative assessment does not occur until the preceding two steps have taken place and the learner is also given feedback to help him determine what to do in order to improve (Wiliam, 2008). Fourth, summative assessment occurs at the end of an instructional episode and does not provide the opportunity for learners to improve their work or understanding. Rather, summative assessment ends with a judgment as to whether the student has mastered the material or not, and instruction proceeds regardless of the results (Popham, 2008; Shavelson et al., 2008).

Of all of these forms of assessment, formative assessment has been shown to result in significant increases in student learning (Black & Wiliam, 1998a; Schroeder, Scott, Tolson, Huang, & Lee, 2007). Furtak cites three “central questions” from the
National Research Council that describe what teachers and students should consider during instruction:

1. Where are you going?
2. Where are you now?
3. How are you going to get there? (Furtak, 2006, p. 4).

These questions form the basis of formative assessment practice. If teachers and students do not clearly understand what their learning goals are and do not have a plan for working toward them, enduring understanding is not likely to result.

Various approaches to formative assessment can be placed along a continuum describing the extent of pre-planning. Informal and unplanned formative assessment usually occurs spontaneously in response to observations of student work during a lesson. Although this immediate response can be very effective, many teachers, particularly novice teachers, find it difficult to adapt their instruction rapidly or to be able to respond appropriately. Formative assessment can also be more planned yet still involve classroom interactions. This might take the form of teacher questioning using pre-planned prompts embedded within class lectures and discussions. The most formal and planned formative assessments are embedded within curricula to check whether students have met certain learning goals before instruction proceeds much farther (Shavelson et al., 2008).

Models of Formative Assessment

With the emerging awareness of the importance of formative assessment in increasing student understanding and achievement, a number of research groups have developed approaches to understanding formative assessment practice. Studies have
been conducted in various nations, geographic settings, school types, and with students of various ages. The studies themselves include quantitative as well as qualitative approaches to the research methodology, and examine the effects of different types of interventions. The following section of this review will describe some of the more recent and/or heavily cited models of formative assessment as they apply to science classrooms.

*The Black and Wiliam model.* Black and Wiliam are adamant that assessments are formative only if they in some way shape the learning of the student or students. In order to accomplish this, the information that the teacher gains during any kind of assessment sequence needs to be interpreted and somehow used to change what might have normally been done in the absence of such information. If an assessment is given, but the teacher continues with classroom activities without regard to the responses, the assessment was not formative (2006).

According to Wiliam (2006), although “the provision of high-quality tools may be a necessary condition, it is certainly not a sufficient condition for the improvement of formative assessment practice” (p. 287). Teachers must also integrate the tools of formative assessment into their regular classroom practice (Wiliam, 2006). For instance, a professional development workshop or class can show teachers how to use a formative assessment tool, but unless the teacher implements it and learns how to use that information to adapt instruction, formative assessment practice is not enhanced. Researchers need to investigate the ways in which the data about student learning are used by teachers. Furthermore, other measures of how well teachers have implemented formative assessment tools have been integrated into their overall classroom practice may shed light on the role of technology in supporting formative assessment practice.
Black and Wiliam and colleagues report four main categories of classroom practice that are essential to rich formative assessment (Black, Harrison, Lee, Marshall, & Wiliam, 2004; Priestley & Sime, 2005). First, questioning is an important skill to develop if formative assessment practice is to be improved. Particularly, the “wait time” between the asking of the question and the student response needs to be sufficient to allow student reflection and processing. Otherwise, the questions deteriorate to requiring solely factual recall. Other strategies that encourage students to consider thoroughly their responses are also beneficial (Black et al., 2004).

The second main area of rich formative assessment practice is the role of giving students feedback on their work. This may take the form of grading the work, but students tend to dismiss other feedback when a grade is assigned. One strategy suggested for increasing the degree of valuable feedback is to have students rework assignments based on the feedback given. The feedback, in any case, should increase the extent and level of students thinking about their own work. If students are not shown what was well-done and encouraged to respond to those comments, the assessment becomes summative in nature and does not lead to increased student learning (Black et al., 2004).

A third focal area of formative assessment should be an emphasis on student self-assessment and peer assessment. Peer assessment can be very valuable in helping students to master concepts, as they may be more likely to listen to peers than to teachers. Instruction in how to accomplish peer and self-assessment will probably be necessary in many classroom situations to make it as rich as possible. Students should be given clear aims and goals, as well as the criteria by which their work will be judged, such as scoring rubrics. The use of peer and self-assessment is also a critical component of the fourth
and final aspect of formative assessment described by these authors. The use of summative assessments in a formative manner can be particularly useful in helping students get ready for testing situations. The results of summative evaluations can be used to help students identify areas of strengths and areas in which they need to improve (Black et al., 2004).

Overall, the main limitation of this model is that, although the authors emphasize that formative assessment is not formative unless it leads to changes in instruction that enhance learning, the aspects of the model do not address this need. Once the teacher has gathered information about students and discovered that they need more work or different approaches to help students understand a certain concept, how is she to implement the new instruction? The use of peer instruction and peer assessment present some notion of a change in instruction, as students rework assignments based on feedback, but this is the only area in which instructional change is even suggested. The role of feedback is focused on improving the quality of work, but not addressing the underlying conceptual understandings that students hold. This model would be enhanced if it incorporated some notion of conceptual change or other method of addressing students’ deep understanding.

*The Torrance and Pryor model.* Torrance and Pryor describe a model for formative assessment grounded in social constructivist theory. Since assessment occurs in a social environment with interaction between teachers and students, individual knowledge is, at the least, partially socially mediated. Further, formative assessment allows teachers to find out where her pupils’ “zones of proximal development” are and tailor instruction accordingly in order to scaffold their understanding (Torrance, 1993; Torrance & Pryor, 1998, 2001).
Torrance and Pryor also describe a critical difference in classroom assessment. *Convergent* types of assessment focus on whether students have mastered conceptual information. It tends to be curriculum-driven, evaluates students based on whether a correct answer was given, and treats students as the passive absorbers of knowledge. This type of assessment emerges from a behaviorist tradition rather than a constructivist perspective, as it measures how well students have mastered the required content. By contrast, *divergent* assessment comes from the social constructivist position, and focuses on what students understand about a concept rather than what knowledge they can recite. The teacher is interested in finding out what misconceptions students hold and how they articulate their learning. Divergent assessment usually varies from the typical assessment cycle consisting of teacher initiation-student response-evaluation from teacher (IRE) or initiation-response-feedback (IRF). Students are given tasks that are more open and are encouraged to practice metacognitive strategies. Because of these differences, convergent assessment is frequently treated as continual and frequent summative assessment on a smaller scale, while divergent assessment holds truer to most researchers’ notions of formative assessment. Divergent assessment, therefore, is considered by the authors of this model to be more powerful (Torrance & Pryor, 1998, 2001).

The Torrance and Pryor model of formative assessment itself encompasses four main aspects. First, task and quality criteria are communicated to, or even generated in conversation with, students. This can occur at the beginning of a lesson, or it can emerge in conversation with individual pupils. Similarly, the teacher might announce pre-planned task and quality criteria to the class, or the criteria might emerge from the teacher
noticing how students are completing work and what the product is. Feedback given to students would comprise the communication of these criteria in this case (Torrance & Pryor, 2001).

Second, teachers collect information about student understanding through careful questioning strategies. This might include the teacher asking probing questions, asking for clarification of a process, or asking students metacognitive-type questions. This stage of formative assessment should be carried out in partnership with the third aspect, observation of students. Torrance and Pryor report that the teachers in the study described earlier found this to be an important step to focus on after initial questioning, particularly in order to find out how students understand the material. Particularly if a teacher is walking around from lab group to lab group, she needs to take time to discover what thought processes have already been occurring, before jumping in to solve students’ difficulties (Torrance & Pryor, 2001).

Finally, feedback to the students about their performance and learning is a critical aspect of formative assessment. Teachers and students alike are accustomed to summary grades for work, and this is a difficult aspect of practice to change. Feedback should reflect the initial task and quality criteria communicated to and developed with students initially, so that students focus on the learning goals that they were supposed to attain. The feedback could be on either the process that students are carrying out, or on the product generated as a result of their work (Torrance & Pryor, 2001). Interestingly, a study by Butler (1988) indicated that students receiving feedback only in the form of comments, without grade markings, performed better on post-tests compared to similar peers who received either grade marking only or a combination of grades and comments.
Adding an alphanumeric grade or score to otherwise rich feedback appears to diminish the role of feedback in improving work quality. When students are presented with a grade, they seem to dismiss the importance of using feedback to improve their work.

The aspects of formative assessment described by Torrance and Pryor (1998; 2001) have been categorized further into sets of specific classroom activities mediated by the teacher and their possible intentions and effects on students. These activities correspond to phases in a formative assessment cycle and serve as points of reference in describing the process of classroom-based formative assessment. Action research participants working with the Torrance and Pryor model found the cyclic formative assessment model to be a useful framework for evaluating their own teaching practice. However, as their own sophistication with formative assessment increased, teacher researchers found that their actual practice moved between various categories or could be described by several categories merged together (Torrance & Pryor, 1998, 2001).

The Cowie and Bell model. Cowie and Bell emphasize that formative assessment (FA) is “the process used by teachers and students to recognize and respond to student learning in order to enhance student learning, during the learning” (Cowie & Bell, 1999, p. 101). Furthermore, formative assessment is only formative if it leads to action on the part of the teacher to enhance student learning in some way (Bell & Cowie, 2001).

Cowie and Bell describe two basic types of formative assessment: *planned* and *interactive*. Common to both of these is the notion that there must be a clear purpose to conducting the assessment. The purposes may be, however, different for the different types of formative assessment. In planned assessment, the purpose is generally to engage
the entire class in identifying progress toward learning goals and is often used to identify areas that students are struggling with so that instruction can be designed accordingly.

By contrast, interactive FA focuses on individual students or groups, and involves assessment of student learning as they are working on specific learning activities. Since interactive FA responds to demonstrated student needs, interactive FA is less curriculum-driven than planned formative assessment, which measures how well students are progressing toward the required understandings (Bell & Cowie, 2001).

Planned formative assessment incorporates three different phases (Bell & Cowie, 2001). First, the teacher needs to *elicit* information from the students. Frequently these assessments are written and conducted in a semi-formal manner; this allows the teacher to save responses and reflect on them for later action. Next, the teacher needs to *interpret* the information. In planned assessment, information about student learning is generally criterion-referenced data, such as determining whether students have met the various learning standards for a particular unit of study or a given lesson. Teachers should be able to use their prior experience here in evaluating what information students really need in order to progress. Finally, planned formative assessment needs to result in an *action*. This requires teachers to have flexibility in their plans; a curriculum-driven environment that requires specific topics to be taught on each given day is not conducive to action. Teachers need to be able to adapt their instruction to students’ needs (Bell & Cowie, 2001).

With interactive formative assessment, the three phases are similar to those for planned FA, but reflect the different purpose of this kind of assessment. First, the teacher has to *notice* what the student understands. If students are working in small groups or
individually, this is often in the form of the teacher circulating among students and overhearing bits of conversation or glancing at student progress on worksheets or manipulatives. This is a faster mechanism than the more formal eliciting phase of planned FA. The teacher also has to recognize the importance of the information she is gathering about students. This requires the teacher to know how the information fits into the curriculum and to know potential misconceptions. She needs to understand the implications of the information she has gathered. Finally, the teacher needs to respond to the information gathered. This phase is similar to the acting phase of planned formative assessment, but occurs on a faster time scale. Generally, the teacher will use the information she recognizes to generate an explanation or demonstration for the individual or group. Often, this explanation or demonstration will then be directed back to the whole class, if she feels that it would be an appropriate clarification for other students who might be struggling with similar concepts (Bell & Cowie, 2001).

Although interactive formative assessment is more immediately responsive to student needs, it does have the drawback that if the teacher is not present at the time the utterance is made by a student, the formative assessment opportunity is lost. It provides more detailed information regarding a subset of students’ understanding, but does not allow for the assessment of all students’ understanding simultaneously. An interesting feature of connected classroom technology is that it allows the more rapid feedback of interactive formative assessment while eliciting evidence of understanding from the whole class at one time.

*The Ruiz-Primo and Furtak model.* This model of formative assessment focuses on the application of ESRU cycles (Elicit question, Student response, Recognition by
Although this model presents clear ways of measuring formative assessment practice in classrooms, it also poses challenges related to its use in examining aspects of formative assessment. A given using strategy and the beginning of another ESRU cycle following an incomplete cycle are not clearly distinguishable. How do the authors decide whether a conversational pattern consists of E-S-R-E-S-R and a cycle of E-S-R-U, where the U is a decision to probe for deeper understanding? In the latter case, the teacher might ask a second eliciting question (such as “How do you know that?”), the student would respond with some rationale, and the teacher would indicate some recognition of the response. The method by which the authors (and any other researcher wishing to use the same model to investigate formative assessment) would treat extended E-S-R-E-S-R-E-S-R cycles within a classroom dialogue needs to be clarified.

Furthermore, short ES or ESR cycles may overlap with one another, making it difficult to distinguish between recognition and elicit steps. This is particularly true when either the recognition or elicitation is implicit, as the authors reported occurring in a number of instances. In addition, this scheme, like that of IRE/F and others, does not
seem to accommodate episodes where a student might initiate a question of his or her
own. Finally, the coding of any discourse patterns in classrooms must remain true to the
essence of the discussion. Categorizing parts of a dialogue between students and teachers
apart from one another runs the risk of diverting the focus of analysis from how a teacher
uses information about student learning to facilitate deeper understanding of concepts.

*Comparison and summary of the models presented.* A number of features can be
used to compare the various models of formative assessment presented here. Each model
has certain aspects that would lend themselves to a specific kind of research study. For
instance, the research paradigm used to frame the methodology of each of the studies that
led to the formation of these models differed. Both Cowie & Bell (1999) and Torrance &
Pryor (2001) used qualitative measures to investigate classroom assessment, and did not
present links to student achievement. Their reports of assessment are also more strictly
interpretive, not having tabulated specific kinds of formative assessment practice for
comparison between teachers. Ruiz-Primo & Furtak (2006; 2007) used both qualitative
measures and quantitative measures. Transcripts were coded, and the degree of FA was
tabulated and used to compare between teachers. Moreover, various student performance
assessments were scored and used to compare the effectiveness of the various teachers in
the study. Black & Wiliam and colleagues (2004) used primarily quantitative measures
to compare the effectiveness of formative assessment practices in the classrooms they
studied.

Differences also exist in the age level and subject area of the classrooms in which
formative assessment practices were investigated. Torrance & Pryor (2001) worked with
teachers in primary grades, with a range of classroom subjects including English, science,
and mathematics. Cowie & Bell (1999) investigated science classrooms at middle school grade levels (ages 11-14). Ruiz-Primo & Furtak (2006; 2007) also worked with teachers in middle school classrooms, within the specific context of the Foundational Approaches in Science Teaching (FAST) program. Black & Wiliam (Wiliam et al., 2004) worked with teachers in middle and high schools, both mathematics and science classrooms, involving students in year 7 through year 11 (approximately equivalent to U.S. grades 6 through 10). Formative assessment practice seems to cut across grade levels and, to a certain extent, across disciplines as well. Science does have a distinct aspect of student understanding, which is that of laboratory exercises, procedures, data manipulation, and the use of experimental evidence. Models of formative assessment that do not take into account the different classroom practices that frequently occur in science classrooms will not be most appropriate for use in a study of middle and high school science classrooms.

The four models described here have varying levels of adherence to a specific cycle of phases that make up formative assessment practice. The model described by Black and Wiliam does not outline a cycle of formative assessment but rather describes certain elements of effective formative assessment (Black et al., 2004; Wiliam et al., 2004). Torrance and Pryor (2001) describe a generalized cycle of formative assessment involving setting task and quality criteria, questioning, observing, and assigning feedback. Teachers in their study emphasized that they experienced fluidity between and among categories as they implemented and reflected on formative assessment practice. Cowie and Bell (1999) describe two different types of formative assessment that occur in science classrooms, each of which has a different set of three phases that comprise a complete formative assessment cycle. Ruiz-Primo and Furtak (2006; 2007) outline the
most detailed set of formative assessment descriptors with their ESRU cycle. Their use of the model in a research study also emphasizes the sequential nature of discrete steps to be taken in conducting formative assessment as they compare teachers on the basis of how many complete and incomplete ESRU cycles are observed.

Clearly, the different models discussed here all have specific aspects that can be used in different contexts and to answer different research questions. The next section of this review will describe a generalized formative assessment cycle derived from the four models presented above. In the view to follow, formative assessment cycles are described as consisting of the administration of an instructional task, the use of quality questioning to probe student understanding, and follow-up actions on the part of the teacher.

*Components of Formative Assessment*

In the context of professional development for mathematics and science teachers, Black, Wiliam, and colleagues found that teachers who developed individualized action plans for improvement in FA practices saw gains in student achievement. Specific strategies that teachers incorporated into their action research plans included increasing and improving questioning of students, using student self-assessment opportunities, feedback in the form of comments with no letter or numerical grades, and making learning goals explicit and visible to students (Wiliam et al., 2004). Although the variety of different measures used for pre-test and post-test achievement measures confound the ability to compare and generalize from the authors’ findings, effect sizes from implementing teacher action plans ranged from -0.35 to 1.55 with a median effect size of 0.27 and a mean of 0.34 (Wiliam et al., 2004). This work does not point to any specific
strategy as being most productive in yielding student gains, but taken as a whole, this study does indicate that the intentional use of FA strategies of whatever type does convey an academic benefit to students as measured by national tests.

Another study on formative assessment used action research by teacher participants to inform an emerging model of FA practices. Harry Torrance and John Pryor (2001) followed up on an earlier study of forty-five elementary grade teachers and administrators in which they used interviews and classroom observations to learn about teachers’ classroom assessment practices (Torrance & Pryor, 1998). Throughout, professional development in the form of half-, whole-, and multi-day workshops was made available to participants for the purpose of instructing teachers in assessment and learning issues, discussion and design of action research, and presentations of progress and emerging data. As part of the intervention, teachers analyzed their own teaching practices with respect to FA, and developed strategies to improve upon their existing pedagogy. Many teachers identified the need for their own teaching to involve more divergent assessment rather than the typical convergent assessment that was so prevalent in their classrooms. They also identified making the purpose of classroom activities and the expectations for quality work, as well as the need to use a variety of questioning and feedback approaches (Torrance & Pryor, 2001).

These two sets of research studies demonstrate that intentional engagement in formative assessment practices can be fruitful in terms of student achievement. The researchers in both of these studies highlight the importance of the teacher clarifying the purpose of instructional tasks, developing stronger questioning skills, and engaging in follow-up to student responses, as described below.
The teacher presents an instructional task. In order to know what students understand about a given topic, teachers need to assign them tasks or question prompts. These can be very formal diagnostic assessments in written form, or might be as informal as the teacher monitoring progress during seatwork or other student activities.

An example of a kind of instructional task is the “formative assessment probe” (Keeley, Eberle, & Tugel, 2007). In this kind of instructional task, students are presented with a scenario or question prompt. A number of choices follow; some prompts use a smaller number of multiple choice responses and others provide the student with twenty or so choices from which the student is asked to select all choices that apply to the question prompt. After the student makes his or her choices, the probes ask the student to provide an explanation for why they chose what they did. Suggested use of these assessment probes includes discussing results as a class so that students can hear why various choices are or are not appropriate. This assessment strategy is formative because the emphasis is not on which answers are correct, but rather how the student thinks about the concepts involved and what meaning the student has made of the question and the concept (Keeley et al., 2007).

Walsh and Sattes (2005) use four attributes to describe quality questions. Quality questions are those that “promote one or more carefully defined instructional purposes, focus on important content, facilitate thinking at a stipulated cognitive level, and communicate clearly what is being asked” (p.23). The first of these attributes relates directly to this notion of an instructional task. According to Walsh and Sattes (2005), purposes might be as wide-ranging as test review questioning, promoting small-group
discussion, comprehension checking, pre-assessment of lesson content, and helping students make connections between their understandings and real-world experiences.

Some of the instructional tasks assigned in a classroom might be pre-planned, but many are spontaneous, drawing on the perceived needs of students. Bell and Cowie found that teachers tended to engage in either planned formative assessment or interactive formative assessment. In planned formative assessment, teachers used semi-formal and written ways of eliciting understanding of the entire class. This was often used at the beginning of a class period to identify which concepts students needed assistance with, as well as where the teacher should start the lesson, conceptually speaking. Planned formative assessment typically had a clear focus on meeting curricular objectives (Cowie & Bell, 1999). Interactive formative assessment, on the other hand, took place generally while students were engaged in small group activities and the teacher would notice particular aspects of student understanding as she moved about the room. In order for teachers to be able to implement interactive formative assessment well, they needed to be able to not only recognize the information as a piece of assessment evidence, but also understand its importance and be able to consider how to assist students in acquiring new information (Cowie & Bell, 1999). Although the purpose for each of these modes of assessment is different, having a purpose to the assessment is a critical component of assessment practice.

Similarly, Torrance and Pryor (2001) describe elements of providing instructional tasks in their description of formative assessment cycles. In their scheme, the teacher must provide students with certain criteria for the completion and accuracy of desired tasks and observe the work both in progress and after completion (Torrance & Pryor,
Ruiz-Primo and Furtak (2006; 2007) also describe the importance of “eliciting questions.” These are a type of instructional task in that they are used to probe student understanding. Overall, the initiation of any formative assessment cycle needs to begin with some kind of activity or question that learners engage in and that teachers can use to collect formal or informal data on student understanding.

The teacher poses questions to elicit student understanding. Earlier in this review, Walsh and Sattes’ (2005) attributes of quality questions were identified. The latter three relate to the kind of question itself: the content of the question, the cognitive level of the question, and the clarity of the question. The content of a question may certainly be thought of in terms of the specific science discipline and topic within that discipline: for instance, a question might relate to photosynthesis. Much of the remainder of this section will describe various aspects of questioning and cognitive levels.

Of the many different ways to assess and classify the cognitive level of questions and responses occurring in a classroom, Bloom’s Taxonomy is perhaps the best-known. Since the original description of this hierarchical categorization of learning objectives, various modifications of this taxonomy have been reported (Anderson et al., 2001; Marzano, 2001). The original Bloom’s Taxonomy identified six levels of instructional objective: Knowledge, Comprehension, Application, Analysis, Evaluation, and Synthesis.

Bloom and colleagues (1956, cited in Anderson & Sosniak, 1994) caution users that when applying the taxonomy to assessment questions, they should consider the context in which the item is assigned. Particularly, teachers need to consider the prior instruction students received as well as the rest of the assessment instrument where the item is located. This is especially important in determining the level of an item which
might on the surface appear to be at a high level, such as analysis, where classroom instruction was such that students are able to memorize the response desired by the teacher, reducing this potentially higher-order question to a recall level.

A recent revision to Bloom’s Taxonomy by Anderson and colleagues (2001) proposes a slight adaptation by changing the category names from nouns to verbs (Remembering, Understanding, Applying, Analyzing, Evaluating, and Creating). Furthermore, this revision has changed the order of synthesis and evaluation. The rationale for this is that true creation from underlying principles is a more difficult task than to make judgments about an item or concept’s properties and utility.

In Bloom’s original taxonomy as well as in the Anderson revision, the categories are hierarchical in that underlying skills and abilities are needed to address fully higher levels of cognition. Skills related to the initial categories of knowledge and comprehension are often needed in order for a student to adequately address higher order questions: they are necessary but not sufficient conditions for being able to carry out, for example, an evaluative task.

Also in Anderson and Krathwohl’s (2001) revision of Bloom’s Taxonomy is a description of multiple dimensions of knowledge. The six levels of instructional objective or assessment item described above are part of a cognitive processing dimension. A second knowledge dimension is described by Anderson and Krathwohl (2001), consisting of four components: factual, conceptual, procedural, and metacognitive. The factual knowledge domain includes things that student need to know about a subject area or concept. The conceptual knowledge domain includes the ways in which other concepts and theories of the subject matter can be linked. Procedural
knowledge includes specific skills that students need to master, such as mathematical algorithms and the “scientific method.” Finally, the metacognitive knowledge domain includes the students’ understandings about how to learn and how to perform at school.

Bloom’s Taxonomy and modifications thereof have been used in a variety of classroom settings for assessment purposes. Eber and Parker (2007) describe Bloom’s taxonomy as a way of providing university-level students in human services programs with the real-world skills they will need to interact with clients. They point out that mere recall of facts, despite being a prerequisite to higher order skills, is of little benefit to a practitioner. Instructors of preparatory courses should therefore provide both experiences and assessments across the full spectrum of question types (Eber & Parker, 2007).

Kasteberg (2003) also works with students in a university setting, instructing college-level algebra. She structures her tests according to the matrix described in Anderson and others (2001), designing items that reflect a variety of question types. Kasteberg (2003) also emphasizes the need to align the instructional materials with the intended assessment strategies.

Although it describes a popular and flexible method of classifying questions and learning objectives, Bloom’s Taxonomy addresses general classroom situations and does not account for specific instructional circumstances. Science classrooms present certain instructional situations that require the use of more specific classification schemes. Some of these schemes address such aspects of science learning as the use of creativity, questions aimed at generating investigative procedures, and questions used during science lectures (reviewed in Chin, 2004). For instance, paired observations among science teachers in an English grammar school resulted in the categorization of questions
including: probing questions, to gather more focused and detailed information; reflective questions, which clarify specific points; and hypothetical questions, which ask students to perform explorations or comparisons with prior knowledge. This classification scheme also distinguished between closed questions, where students responded with just a few words to demonstrate understanding, and open questions, which prompted students to provide their thoughts and ideas without constraint (Carr, 1998; Chin, 2004). The notion of open and closed questions are akin to Torrance and Pryor’s (2001) descriptions of divergent and convergent questions respectively.

Ruiz-Primo and Furtak (2006; 2007) offer a different categorization of questions. They evaluate formative assessment cycles regarding epistemic, conceptual, and social structures. Epistemic aspects of formative assessment are those that concern how students know what they know, how they have interpreted laboratory evidence, and what procedures they may need to follow to complete a task. This is a particularly important aspect of formative assessment in a science classroom, where much of the instruction (ideally) revolves around laboratory investigations. Conceptual aspects of formative assessment relate to the knowledge and understanding of the information that underlies scientific concepts. This would include basic comprehension and knowledge. Finally, social aspects of formative assessment involve the ways that students communicate and represent information. This can include verbal means, such as the formation of coherent arguments, as well as written and other visual displays of information. In their studies of classroom-based formative assessment, the authors classified parts of ESRU assessment cycles as belonging to epistemic or conceptual domains (Ruiz-Primo & Furtak, 2006, 2007). Interestingly, the conceptual domains appear to correlate well with Bloom’s
Taxonomy lower levels of cognition (knowledge and comprehension, with some application of basic concepts) while the epistemic domains require students to function at higher cognitive levels (application, analysis, and evaluation of information).

Questioning strategies have been an important aspect of the educational process since the work of Socrates became known through Plato’s *Dialogues*. By asking questions, Socrates and his pupils gradually come to acquire understanding of the world (Rowe, 1974). As the answers to these questions are not known but prompt further inquiry on the part of the student and the teacher, today’s education professionals would classify the so-called Socratic Method as a means of engaging in higher-order questioning strategies.

Critical thinking skills are demanded both in the classroom environment as well as in the workplace. Educators are in wide agreement that the development of such skills is a critical aspect of K-12 education. Further, formal statewide assessments rely on the use of higher order questioning techniques, requiring that students be competent at interpreting and fully responding to questions at a variety of cognitive levels. Many textbooks are following this trend by ensuring that “Section Review” and “End of Chapter Questions,” as well as the more formal chapter and unit assessments, utilize items of a variety of cognitive levels.

Students construct meaning from pre-existing knowledge (National Research Council, 2000). Referred to as a constructivist viewpoint, this notion requires that students’ prior knowledge is determined, and that student perceptions are monitored at all stages during the learning process. Quality questioning strategies are an integral component of this process. Scientific study, in particular, requires a variety of higher-
order thinking skills, such as the ability to hypothesize, predict, design experiments, and analyze data (National Research Council, 1996). To ensure proper development of these skills, students must be consistently engaged in activities and cognitive tasks that rely on the use of higher order thinking skills.

Prior studies have examined the degree of teacher questioning and the level of questioning. One such study used a computer input method to record which of four basic types of question a teacher used at any given time period in social studies classes of a wide range of grade level. Question types identified included:

- literal questions, which are similar to Bloom’s Knowledge level and require recall of facts and meaning;
- interpretive questions, which are similar to Bloom’s Comprehension level and ask students provide an interpretation or explanation;
- application, which require students to apply knowledge to a new situation and could be considered similar to Bloom’s application level; and
- affective questions, which include all cases of students being asked to consider their own values, interests, beliefs, or attitudes (Daines, 2001).

While 40% of the class time was found to be spent in teacher-directed questioning, the majority of these questions were at the literal recall level rather than higher order thinking such as interpretation and application, or in reflective questioning addressing the affective domain (Daines, 2001). This study examined only three cognitive levels of questions and focused on measuring frequency of responses.

Other studies involving higher order questioning in science classrooms focus on specific critical-thinking instruction. In one particular study, middle-school students
were provided with instruction in developing quality investigative questions. Results of a pretest/posttest protocol indicate that the intervention had a small effect on students’ ability to develop researchable questions (Cuccio-Schirripa & Steiner, 2000). Another research report used case studies to engage students in problem-solving modules. In this study, students in experimental groups increased their reasoning skills over students not receiving critical thinking instruction and practice (Zohar & Dori, 2003).

To increase students’ abilities to engage in critical thinking and respond to higher-order questions, students must be exposed to rich questioning practice on a regular basis. However, despite many efforts of science education reform, whole-group direct instruction with sporadic, loosely connected hands-on and laboratory exercises remains the primary mode of instruction in many science classrooms. A critical area of education research, then, is elucidating instructional strategies that support rich questioning practice.

The teacher follows up on student responses. The meaning of follow-up used in this study is that of a teacher providing comments to individual students or classes as a whole, based on the evidence the teacher has about student learning. Many researchers agree that the feedback component of grading is critical to the potential increases in student understanding. Black and Wiliam point out that quality feedback does not include grades but rather constructive, qualitative comments that encourage students to re-examine their own work and seek to improve it (Black et al., 2004).

Not all studies of feedback show increases in student learning. In a review of formative assessment studies, Baroudi (2007) points out that effective feedback focuses on the product rather than the student. This is echoed by a feedback classification system
that helps teachers identify effective and less-effective forms of feedback (Brookhart, 2008). Feedback, according to Brookhart (2008), should be directed at the work the student completed or the strategies and procedures the student used to accomplish the designated task.

In another extensive review of the literature on feedback, Shute (2008) summarizes findings of many studies to provide guidelines for effective feedback, grouped according to feedback related to enhanced learning, timing, and characteristics of learners. Identified in this review is also the notion of depersonalizing feedback. Additional recommendations include considering the medium (written or oral) and the setting (personal, small-group, or whole-class) in which feedback is administered. Other practices that were shown by the reviewed articles to interfere with effective feedback practices include not interrupting an active learner to issue feedback and limiting praise (Shute, 2008).

Feedback has been demonstrated to be a major component of formative assessment by a number of research groups. In studies of teachers conducting action research to select and implement strategies that support their formative assessment, Wiliam and colleagues identified certain features that benefit student learning as well as teacher learning about students. Specific strategies that teachers incorporated into their action research plans included increasing and improving questioning of students, using student self-assessment opportunities, feedback in the form of comments with no letter or numerical grades, and making learning goals explicit and visible to students (Wiliam et al., 2004). In other studies by Bell and Cowie (2001; Cowie & Bell, 1999), the intentional use of formative assessment is not complete unless some action is taken by the
teacher to respond to the needs of the students. Similarly, Torrance and Pryor (1998; 2001) indicate evaluative feedback as being an essential component of their framework for formative assessment. Finally, Ruiz-Primo and Furtak (2006; 2007) use the ESRU cycle to highlight the role of teacher responses and uses of the gathered information.

Formative assessment, then, incorporates four main aspects. Teachers implement instructional tasks, question students to uncover understanding, become aware of student learning, and follow-up on what is learned about a given concept. The relationship of these four phases of formative assessment and their hypothesized effect on student learning is described in Figure 2.1. In the top row of Figure 2.1, instructional tasks that do not encourage rich student thinking are coupled with weak questioning strategies. This leads to the teacher’s lack of awareness of student understanding. Without evidence of student understanding, the teacher cannot adapt instruction accordingly, and student learning is not enhanced. This is the situation when formative assessment is not implemented. The second row of Figure 2.1 depicts the situation when formative assessment is implemented but the teacher lacks the necessary skills and/or knowledge to adapt instruction. Although rich instructional tasks and strong questioning skills lead to heightened awareness of student understanding, the teacher is ill-equipped to implement alternate instructional strategies, and student learning is not enhanced.
The final row of Figure 2.1 demonstrates the hypothesized effect of rich formative assessment on student learning. High-quality instructional tasks are combined with thorough probing of student understanding in order that the teacher know what students know. When coupled with the pedagogical skill and content knowledge necessary to develop alternative instructional strategies to address misunderstandings, teacher knowledge of student understanding may result in increased student learning. In the next section of this literature review, pedagogical and technological strategies that support these aspects of formative assessment will be described.
Interactive Classroom-Based Formative Assessment

The Interactive Classroom

Students who are actively engaged in their own learning process are likely to learn more and have deeper understanding. In an interactive classroom, students are encouraged to ask questions, respond to the teacher, engage in discussion with peers, and reflect on their own learning (Ridley, 2007). Interactive teaching methods include the use of cooperative learning groups, problem-based learning, and other student-centered instructional strategies. Ridley (2007) cites a number of studies that demonstrate a link between the use of interactive teaching strategies and student learning. Active learning has been demonstrated to positively affect students’ problem solving and critical thinking skills (Sivan, Leung, Woon, & Kember, 2000) and student learning (Knight & Wood, 2005). Since the degree of interactivity appears to influence student learning in these classrooms, how would a teacher create a classroom environment that supports active learning?

How People Learn (National Research Council, 2000) specifies four types of “centeredness” that promote interactive learning and rich instructional practices. First, the instructional environment can be learner-centered. Learner-centered instruction recognizes that the classroom environment consists of individuals from varying backgrounds and perspectives, and that students bring prior understandings into this environment. Learner-centered teachers “attempt to get a sense of what each student knows and can do as well as their interests and passions” (National Research Council, 2000, p. 136).
Second, interactive learning environments are knowledge-centered (National Research Council, 2000). Obviously, specific concepts and skills must be mastered in the study of any subject area. However, knowledge-centered environments go beyond knowledge and skill acquisition. In a knowledge-centered classroom, the teacher works with the students to help them make sense of the information. Connections among various concepts and disciplines are developed that help students to understand not only the information they encounter but also how it relates to themselves, other areas of study, and the rest of the world (National Research Council, 2000).

A third characteristic of interactive classrooms is that they are assessment-centered (National Research Council, 2000). Teachers continually monitor students for understanding, ensuring that learning goals are being addressed and providing feedback to students to help them improve their work. A variety of styles of assessment, including performance-based, authentic, and more traditional forms of assessment help the teacher, the student, and other stakeholders to understand what the student knows and can do (National Research Council, 2000).

Finally, a fourth type of centeredness encompasses all of these three mentioned above. Community centeredness is the notion that the classroom norms and expectations are established in such a way as to honor the contributions of all students. Community centered classrooms also include the various relationships between the school and the external community, such as homes, businesses, and other community institutions. The majority of a student’s time is spent outside of the classroom; a strong teacher will not ignore the contributions of outside influences on the learning process (National Research Council, 2000).
Various non-computer-based pedagogical techniques have been developed to increase the degree of interactivity in classrooms. For instance, the use of response cards to promote formative assessment meets the centeredness criteria as mentioned above. With response cards, students are asked to hold up a card that indicates their response to a given prompt. Several studies of response-card use in college-level classrooms demonstrate increased student participation, increased student engagement, and increased student achievement on course assessments (Kellum, Carr, & Dozier, 2001; Marmolejo, Wilder, & Bradley, 2004; Preszler, Dawe, Shuster, & Shuster, 2007; Randolph, 2007). Similar results were seen in elementary school settings in science or social studies classrooms using response cards (Gardner III, Heward, & Grossi, 1994; Narayan, Heward, Gardner III, Courson, & Omness, 1990).

If using response cards promotes student learning, why should teachers investigate using more technologically sophisticated tools? After all, they are much more expensive and time-consuming to set up than response cards or even having students raise their hands to indicate responses. Abrahamson (2005) suggests several reasons why the use of interactive technology to promote formative assessment can bypass certain difficulties with other methods. Many students value anonymity in their responses: perhaps they are afraid of giving a very public incorrect response. With hand raising or response card flashing, the rest of the class can readily see what each student has indicated. This leads to a secondary issue, that of students piggy-backing on others’ responses. A slow responder can wait to see what the majority is responding, and then display his or her answer when it seems safe to do so. This replaces cognition that potentially occurs while students select their responses with copy-catting. Another
concern with hand-raising and response cards is the management of a classroom full of data. Teachers may have difficulty quickly aggregating the data (in the form of colored, numbered, or lettered cards or in the form of raised hands) to know whether to proceed with instruction. This problem is exacerbated with large university lecture halls. Additionally, unless the teacher takes the time to mark individual answers, no record of student responses exists. This limits the student being held accountable for his or her response, and can therefore also limit the extent of cognition that goes into selecting a response (Abrahamson, 2005). Other issues with hand raising and response cards are the limited number and types of responses. A teacher is essentially limited to asking questions with a binary choice or at the most, with a small number of multiple-choice responses. Open-ended responses are not practical with these systems of response.

Active learning strategies have been shown to increase student learning. Coupled with classroom environments that are learner-centered, knowledge-centered, assessment-centered, and community-centered, these strategies could have even greater effects on student achievement. One way to increase the “centerednesses” in a classroom is to prompt students to respond to frequent questioning with hand-raising or by holding up response cards. Although these techniques involve students in their own learning, a number of technological assessment systems have the potential to support active learning at even greater levels.

Audience Response Systems

A number of instructional technology innovations have become increasingly common in classrooms from elementary schools through graduate-level work. When used in pedagogically advanced ways, interactive classroom technology supports all four
centerednesses as described above. By promoting the more frequent use of formative assessment resulting in more accurate collection of data regarding student understanding, teachers have a better sense of what students do or do not understand. Connections among concepts can be made more readily. The focus of the classroom shifts from having a teacher lecture to having students engage in dialogue with one another and with the instructor. A sense of classroom community is developed that recognizes and honors the contribution of all students.

Sometimes referred to as audience response systems (ARS), interactive classroom technologies encompass a wide variety of specific hardware technologies. Many university lecture halls are being fitted for wireless communicators called “clickers.” A clicker system generally consists of an IR (infrared) or RF (radio frequency) network to connect a central computer with a number of small handheld devices. The devices, usually smaller than a deck of playing cards, consist of a wireless transmitter and up to ten buttons, depending on the model. To use a clicker system, the instructor activates the clickers, poses a question, and provides time for students to respond. Various forms of presentation software, course management systems, or proprietary display software are used to provide students with the question prompt and reveal aggregated responses. Students use the handheld devices to submit a response, and responses are tabulated and stored. With many systems, the teacher and students can then quickly view a graph depicting how many students chose which response (Burnstein & Lederman, 2001; Burnstein & Lederman, 2003; Caldwell, 2007; Skiba, 2006).

Skiba (2006) describes the benefits of using a clicker system as encouraging responses from all students without fear of being incorrect, rapidly aggregating data
involving multiple choice responses, increasing dialogue between student and teacher or among students, and promoting active learning in the classroom. Further, teachers are equipped to provide better feedback regarding student learning. Ribbens (2007) describes similar experiences, citing his use of aggregated student responses to make rapid decisions whether to reteach or move on to a new topic and the ways in which he facilitated open discussion and collaboration.

Most studies of clicker systems have been in undergraduate or professional school lecture halls. Many of these studies have focused on affective aspects of instruction. For instance, Medina and colleagues (Medina et al., 2008) used clickers in pharmacy classes on two different campuses simultaneously. The instructors faced the challenge of instruction in two locations physically removed from one another. Clicker technology was investigated as a potential way to increase student engagement, particularly in the remote classroom. The authors found that the students in both settings did appreciate the active learning that they attributed to the use of the clicker system, and preferred uses of the clickers that did result in graded assessments. The instructors reported that they found knowing more about student understanding to be very useful (Medina et al., 2008).

Similar results were found in a study of twelve university courses using clicker technology for the first time (Graham, Tripp, Seawright, & Joeckel, 2007). In this study, students were surveyed with a variety of measures to determine their attitudes and beliefs regarding the use of the audience response system. Students preferred uses of the clickers that led to their own self-assessment and the instructor learning more about the class, as opposed to the clickers being used for grading and attendance purposes. The surveys also revealed that students were concerned about the cost of the clickers and
whether they would be worth the cost, considering the amount of use. Some students also indicated that they felt that instructors would use the system as a way to avoid the “busy work” of grading quizzes or to incorporate attendance grades as a way to punish students (Graham et al., 2007).

While these studies highlighted the positive benefits (and some potential concerns) of using audience response systems in large lecture halls, they did not address the issue that is central to formative assessment: did the students participating in the technological innovation learn more? Various other studies have examined clicker technologies by measuring student progress on periodic quizzes and through the use of common final exams. In undergraduate business classes, a quasi-experimental study comparing the immediate feedback of quizzes scored with clickers to traditional quizzes graded and returned the following week showed a significant increase in student achievement (Yourstone, Kraye, & Albaum, 2008). However, it should be noted that the intervention here was not just the use of the clicker technology but also class discussions immediately following quiz scoring. The relative contributions of students seeing how they compare to their peers and discussion about misconceptions and misunderstandings to the increase in student achievement are not clearly understood (Yourstone et al., 2008).

Lasry (2008) also conducted a direct comparison between two undergraduate introductory physics classes at a community college, one using clickers and the other using cardboard flashcards to indicate their responses. He saw no significant difference in the student achievement gains when comparing pre-test and post-test scores. Lasry attributes this absence of a difference in treatment to the ineffectiveness of a technological innovation (Lasry, 2008). However, several aspects of this study present
challenges to the conclusions. First of all, the finding of no significant difference is based on a gain score, the calculation of which is not explicitly described. It appears to be based on the overall class averages on pre-tests and post-tests rather than individual paired student difference scores. Second, the intervention in these classes is not just the presence of cardboard flashcards versus electronic clickers; both types of classes utilized a strategy described by Mazur (1997) as “peer instruction.” In peer instruction, direct lecture-based instruction is interspersed with specific task prompts for students. They consider their responses individually then attempt to convince a neighbor, followed by an opportunity to revise their initial responses. After polling the class (show of hands) additional discussion takes place as needed (Mazur, 1997). In Lasry’s (2008) study, this pedagogical technique is the norm. His classes consist of about 40 students, allowing the visual aggregation data in the form of raised hands or flashcards.

By contrast to these findings, Crossgrove and Curran (2008) found a significant difference in the achievement scores of non-majors enrolled in an introductory college biology course using clickers compared to a non-clicker class. Retention of biology content after four months post-instruction was also significantly different compared to classes with no clickers used. A similar but lesser effect was seen in a second year genetics course for majors. More specifically, the increases in student achievement were found to occur across the major cognitive domains of Bloom’s Taxonomy. The authors do also suggest that some of the difference may be due to the active learning strategies that were used in the courses (Crossgrove & Curran, 2008).

“Clicker”-style audience response systems have positive effects in higher-education classrooms. These benefits are seen both in terms of student achievement as
Interactive Networked Classroom Assessment Systems

Clearly, audience response systems have a tremendous potential benefit in terms of student achievement, student engagement, and ability of the instructor to learn more about what students understand. The studies mentioned above used a type of audience response system that is limited to certain kinds of question stems, primarily multiple choice. Other forms of interactive classroom technology exist, such as the TI-Navigator™ that allow the teacher to design different forms of instructional tasks. This system uses graphing calculators commonly used by students in the sciences and mathematics, and has the added capability of being able to display graphs and equations on the screen. The instructor can send or receive data of a variety of types, in addition to the questions and responses sent through a clicker system. The instructor can rapidly display student work on a computer screen or project it for the entire class. One drawback to this system is that it accommodates only thirty-two calculators on a single system. Further, either the school needs to purchase a class set or each individual student needs to invest in a graphing calculator, at approximately $120 each.

In a quasi-experimental study of middle school students in Texas, algebra I students were shown to experience significant gains in achievement scores in classrooms equipped with the TI-Navigator™ system (Stroup et al., 2005). The authors of this study describe a major advantage to the use of interactive classroom technology being the
ability to elicit from students a wide range of appropriate responses rather than just one correct answer. For example, rather than having students simplify the expression $2x + 3x$, the authors propose giving the prompt “create functions that are the same as $f(x) = 5x$” (Stroup et al., 2005, p. 4). A wide variety of responses could then be received by the teacher and displayed for the class, facilitating a rich discussion about functions and algebraic manipulations.

Classroom Connectivity in Promoting Mathematics and Science Achievement (CCMS) is a large-scale national randomized cross-over trial of over 100 Algebra I and physical science classrooms that have been equipped with the TI-Navigator™ system. Along with the inclusion of the technology into classrooms, the intervention in this study includes a week-long workshop on using the Navigator™ system and the associated pedagogy of a connected classroom as well as a follow-up professional development day and participation at the annual Teachers Teaching with Technology International Conference. A wide variety of measures are used to investigate the experiences of students and teachers during this study, including classroom observations, teacher telephone interviews, student focus groups, student pre- and post- achievement tests, and a number of teacher and student surveys (Pape et al., 2008). Implementation of connected classroom technology is hypothesized to increase the amount and/or effectiveness of formative assessment in mathematics and science classrooms.

Early findings from the first year of the CCMS study show that students enrolled in Algebra I classes where this interactive networked classroom system was in use outperformed their counterparts who had Algebra I classes without this technology with an effect size of 0.37. When other variables, such as school SES, teacher gender, and
teacher years of experience, are accounted for, the difference remains statistically significant, with an effect size of 0.30 (Pape et al., 2008). Continuing investigations with this research project involve examining how student achievement and other characteristics may change when the teacher has more experience using the connected classroom technology. Notably, with this research study, teachers are not issued specific lesson plans or requirements for implementation. Rather, the research team is continuing to investigate how teachers use the technology to promote formative assessment.

Summary

Formative assessment is one of the instructional strategies that has been shown to have a large impact on student learning (Black et al., 2004). Formative assessment involves cycles of the teacher posing instructional tasks to students, eliciting responses and providing students with feedback, and changing instruction either on an individual or class-wide basis. Despite being acknowledged as a promising research-based best-practice for increased student learning, formative assessment has also been shown to be a relatively weak aspect of pedagogy. Teachers may be constrained by external pressures of accountability for testing and curriculum or they may feel ill-equipped to carry out effective formative assessment practices (Daws & Singh, 1996). Teachers need more tools to assist them in incorporating formative assessment into their classrooms, but the presence of a tool in itself does not necessarily translate into more sophisticated formative assessment practice. Rather, the tools must be fully integrated into the teacher’s classroom practice (Wiliam, 2006).
Others have indicated a role for using interactive classroom technologies in promoting formative assessment (Penuel, Tatar, & Roschelle, 2004). Formative assessment can be difficult to carry out for a number of reasons, including the challenge of collecting and aggregating accurate data on student learning in a short timeframe. Handheld devices provide a way to facilitate rapid data collection so that teachers may quickly collect information regarding student understanding. Interactive classroom technology promotes the acquisition of more accurate assessment data, which is what Black and Wiliam (1998b) describe as being one of the keys to successful formative assessment and increased student learning. Formative assessment and interactive classroom technology have been shown to positively affect student learning, both in terms of achievement across content areas as well as in the affective considerations of the learning process.

A major component of classroom-based formative assessment, regardless of the degree of technology support employed, is the role of teacher questioning and the resulting classroom discourse. Several researchers emphasize that increased discussion occurs among students as they respond to question prompts during lectures (Beatty, Gerace, Leonard, & Dufresne, 2006; Ribbens, 2007; Skiba, 2006; Yourstone et al., 2008). Interactive classroom technology allows the teacher to synthesize rapidly whole-class information regarding student understanding in response to specific questioning probes. Teachers engaging in planned FA can input questions in advance or develop prompts on the spot and deliver them to students remotely during class. Results can be displayed for all of the students to see, promoting the opportunity for rich discussion about “right” and “wrong” answers. These are things that are hypothesized to influence formative
assessment practices in connected classrooms. However, a detailed analysis of the ways in which formative assessment is practiced in connected classrooms has not been completed previously.

While collecting and aggregating evidence of student understanding are critical components of the formative assessment process, one must also consider the quality of the instructional tasks and questioning strategies, as well as the extent of the subsequent classroom discussions, in examining formative assessment practice. Technologies that include student interfaces with additional functions of alphanumeric data entry, graphing, and navigating one’s cursor around a screen have even more potential for supporting rich instructional tasks. They also provide different representations of understanding to serve as the basis for classroom discussion.

Although much work has been performed regarding student achievement and affective responses to audience response systems, little is known about the actual process of formative assessment as carried out in these classrooms. Much of the work related to formative assessment in interactive technology-equipped settings examines the role of clicker tools in university lecture halls, rather than secondary science classrooms. The study being described here adds to the body of work in this area by providing a detailed analysis of the teacher and student interactions during formative assessment episodes mediated through interactive networked classroom technology in secondary science classrooms. This contributes a new methodology for studying assessment episodes and questioning cycles. This dissertation study also provides a new model for formative assessment based on the detailed analysis of assessment episodes.
Classroom-based interactive technology is becoming part of a growing number of classrooms. By identifying patterns of formative assessment conducted in these environments, this study will assist educators and researchers in better understanding how teachers carry out formative assessment practice in different contexts. The work presented in this dissertation will also provide additional information regarding how teachers can implement formative assessment strategies with innovative networked assessment tools.
Chapter 3 : Research Methodology - Intervention

The study presented in this dissertation uses data collected as part of a larger research study at The Ohio State University. This chapter describes the intervention of that larger research study. Chapter 4 of this dissertation presents the methodology specific to the guiding question being addressed in this dissertation.

Classroom Connectivity in Mathematics and Science Project

*The Intervention*

Classroom Connectivity in Promoting Mathematics and Science Achievement (CCMS) is a multi-year longitudinal study of the effect of connected classroom technology in over 100 secondary mathematics and science classrooms in 28 different states, and is generously supported, through Grant Number R305K050045 to The Ohio State University, by the Institute for Educational Sciences (IES) division of the U.S. Department of Education. This project utilizes a randomized cross-over design where half of the participants (teachers) receive the intervention in the first year of the study while a second cohort of participants serves as a control group in the first year. In the second year, the control cohort receives the intervention. Thus, the second cohort of participants serves as a control for the first cohort, as well as a control for its own progress. A variety of qualitative and quantitative measures are used to examine
pedagogical aspects of classrooms equipped with a graphing calculator-based audience response system. Teacher implementation and teaching strategies as well as student perceptions and achievement are key variables to be examined in the CCMS study. The intervention for CCMS occurs continually through a teacher’s participation in the project. The project is implemented as a full-scale randomized controlled trial for Algebra I classrooms; additional details regarding the CCMS research study can be found in Pape et al. (2008). In the CCMS project, Physical Science classrooms comprise a pilot study sample. The remainder of this chapter will focus on the intervention as applied to physical science teachers who experienced the project as part of the treatment group in their first year of participation.

For the first phase of the intervention, 10 physical science teachers attended a one-week training course held at The Ohio State University in July 2006. The content of the course was how to use the Texas Instruments (TI) Navigator™ in conjunction with implementing relevant pedagogy, supplemented by examples of teaching physical science with the system. The course was designed and administered by the CCMS project research team in conjunction with workshop leaders who teach physical science in secondary schools. The workshop leaders all implemented graphing calculator technology as well as the TI-Navigator™ system in their regular teaching, which lends authenticity to the workshop instruction. Participants were provided with a workshop notebook containing readings from scholarly and practitioner journals regarding the role of connected classroom technology (CCT) in teaching, self-regulated learning, and formative assessment. The workshop notebook also included extensive systematic instructions on how to set up and use each component of the TI-Navigator™ system. The
workshop was taught in a hands-on and experiential manner, with participants following along with the workshop instructors at their own laptop computers. Time was provided for participants to work independently or in smaller groups to master aspects of the system, and the course culminated with the presentation of participant-designed lessons that they could go back to their schools and implement in the classroom. As part of the research project, school districts where participating teachers worked were given a substantive discount on the purchase of one TI-Navigator™ system for their classroom use.

Figure 3.1 provides an example of a classroom setting including the TI-Navigator™ system. Briefly, each student uses a handheld graphing calculator, often a member of the IT-83 Plus or TI-84 Plus family of calculators. Each calculator is connected by a link cable to a hub. Each hub can accommodate four calculators, and communicates wirelessly with an access point. The access point plugs into the teacher’s computer via a USB port. It receives wireless signals from the hubs and relays the information to the teacher’s computer. The system is most effective when coupled with a digital projector so that all students can see the class’s responses as they are collected on the teacher’s computer screen. Particularly in science classrooms, teachers can pair calculators with scientific probes so that students can measure values such as temperature, light intensity, pH, or elements of motion (distance and time).
In the first year of science participation (February 2007), 9 of the 10 science teachers participating in the CCMS study attended the annual Texas Instruments Teachers Teaching with Technology (T³) International Conference, held in Chicago. The CCMS project used the day before the conference to provide professional development for all research participants attending. This allowed the science cohort members the opportunity to interact with and learn from their colleagues in the mathematics portion of the study. The professional development day included time for technology troubleshooting assistance, working with peers to master additional aspects of the TI-Navigator™ system, and discussions of self-regulated learning and formative assessment.
Teachers were asked to bring with them, on a flash drive, materials for a lesson they had taught using the TI-Navigator™. At the end of the day, the contributed lessons were compiled and distributed to each teacher. Following the professional development day, teachers attended the two-and-one-half day T³ Conference.

Another aspect of the intervention was the regular contact made with teachers by members of the research team. In late autumn, all science teachers from the summer workshop were contacted by telephone to engage in a brief (20-30 minutes) interview (details of the interview are below). This served to check on any emerging issues regarding availability and set-up of the technology, as well as to remind teachers to obtain student consent and to administer student measures. Toward the end of the school year, teachers were contacted again by members of the research team to engage in a slightly longer (30-45 minutes) telephone interview (see details below). The purpose of this interview was to learn more about how the teachers had implemented the technology and pedagogy of connected classrooms and to find out about any plans the teachers had for their ongoing development.

As a final aspect of the intervention and data collection, some teachers were selected for two-day classroom observations. For science participants, all teachers who had implemented the technology at all by the time of the fall interview were chosen for classroom observations. Members of the research team videotaped one or more class sessions on two successive class days. For each observation, the teacher was asked to participate in a post-observation interview regarding the lesson that was videotaped. A small sample of students was chosen to participate in a student focus group to provide additional information on the technological and pedagogical aspects of CCT.
implementation. Table 3.1 provides an overview of the intervention and data collection for science teachers. Details of the professional development and data collection are provided for the first year of physical science implementation, 2006-2007. For teachers continuing for a second or third year of the CCMS study, data collection proceeded as described for the first year of implementation.

Table 3.1. Timeline of Intervention and Data Collection.

<table>
<thead>
<tr>
<th>Time</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>July 2006</td>
<td>Participants attended Summer Institute in Columbus, OH</td>
</tr>
<tr>
<td></td>
<td>Participants completed initial demographic surveys</td>
</tr>
<tr>
<td>September/October 2006</td>
<td>Participants completed teacher measures</td>
</tr>
<tr>
<td></td>
<td>Participants obtained student assent/parent consent</td>
</tr>
<tr>
<td></td>
<td>Participants administered student “pre” measures</td>
</tr>
<tr>
<td>November/December 2006</td>
<td>Fall telephone interviews conducted</td>
</tr>
<tr>
<td>February 2007</td>
<td>Professional Development Day at T³ International Conference, Chicago IL</td>
</tr>
<tr>
<td>February-May 2007</td>
<td>Two-day classroom observations for selected participants</td>
</tr>
<tr>
<td>May 2007</td>
<td>Participants administered student “post” measures</td>
</tr>
<tr>
<td></td>
<td>Participants completed teacher measures</td>
</tr>
<tr>
<td>May/June 2007</td>
<td>Spring telephone interviews</td>
</tr>
<tr>
<td>September 2007-June 2008</td>
<td>Data collection as described for year 1</td>
</tr>
<tr>
<td></td>
<td>Professional Development Day at T³ International Conference in Dallas</td>
</tr>
<tr>
<td>September 2008-June 2009</td>
<td>Data collection as described for year 1</td>
</tr>
<tr>
<td></td>
<td>Professional Development Day at T³ International Conference in Seattle</td>
</tr>
</tbody>
</table>
In their second (2007-2008) and third (2008-2009) years of participation in the research study, science teachers experienced the intervention essentially as described above, except that these participants did not attend a summer workshop. In 2008, the professional development day was held the day before the T³ conference in Dallas and aimed to elicit advice teachers would give to new users of the technology. A final professional development day took place in Seattle in 2009. Ongoing professional development continued through the years of the teacher’s participation in the CCMS project by way of online help through the TI website and a discussion forum of CCMS project participants in the form of an email list-serve.

Quantitative Measures

A wide variety of quantitative measures were implemented in the CCMS research study. In this section, a brief overview of the content of these quantitative measures is provided. Copies of those measures used in the remainder of this dissertation are included as Appendices.

Teacher demographics survey. This survey was administered to participants at the beginning of their week-long summer workshop. Teachers were asked to provide information about their academic background (degrees and licensures earned) and the location of their schools. The demographic survey used for science participants is located in Appendix A.

Demographic data describing the school context in which each participant taught were also collected. Geographic classification of the school setting was obtained from the United States Census Bureau online. Data regarding school-level demographic characteristics were obtained from GreatSchools.net, Schooltree.org, and the National
Center for Education Statistics whenever available. Additional data were obtained by
calling schools to request the information from administrative offices. The percent of
students qualifying for free or reduced lunch benefits at the school level was used as a
proxy for the percent of students from lower socioeconomic backgrounds.

Teacher technology use and professional development survey. As participants
began the week-long summer workshop, they were asked to complete a technology use
and professional development survey. This survey consisted of 85 Likert-type items
asking teachers to describe their comfort with various computer and educational
technologies. Additional items asked teachers how frequently they engaged in
professional development. This survey, located in Appendix B, was completed only once
during the study.

Teacher instructional beliefs and practices survey. The teacher instructional
beliefs and practices survey (TIPBS) was administered online to science teacher
participants in the fall of their first year of participation in the project and at yearly
subsequent time points. The first administration of this survey consisted of 195 items,
with most being scored on Likert-type scales. Some open-ended items were also
included for teachers to describe their classroom and school populations. A modified
version was administered in the spring of their first year of participation. This shorter
version consisted of 132 items, also primarily Likert-style, with some open ended items
regarding curriculum and textbook selection.

Student perceptions of instruction in science. This measure was administered to
students only once per year, at the conclusion of the course in which Navigator™ was
used. It consists of 39 Likert-style items designed to collect information about the classroom environment and teaching practices from the student’s perspective.

**Student achievement in physical science.** The preparation of the achievement test for physical science has been described elsewhere (Shirley, Irving, Sanalan, Boscardin, & Ucar, 2006). Briefly, secondary science standards from the states of greatest population were tabulated to generate a list of the most common benchmarks. For most of these standards, released items from state graduate and end-of-course exams were used with permission. In a few cases, items addressing the standard were written by the research team. An instrument consisting of 46 items was piloted in two different school districts, with 269 ninth grade students participating. Items with extreme difficulty or low discrimination were identified through item response theory; four items were removed as a result. The resulting 42-item instrument was administered to all students as a pre-test measure.

For most teachers, this instrument was used intact as a posttest, five to eight months later. Since the target audience for this test is 8th and 9th grade physical science classes, some classes in this study cover more advanced content than is addressed in this test, and students are likely to have previously taken an introductory physical science course. For example, participants may teach upper-level Chemistry or Physics courses, rather than the more general physical science course. For upper-level physics classes, the pretest was modified by removing non-physics related items and replacing them with items from the Force Concept Inventory (Hestenes, Wells, & Swackhamer, 1992).
Student papers were mailed by the participating teacher to a scoring facility. A database of student scores was sent electronically to the research team, who then cleaned the data to ensure than no student without consent remained in the database.

*Qualitative Measures*

*Telephone interviews.* Fall telephone interviews were conducted between November and January of the 2007-2008 academic year. One researcher conducted all science teacher fall interviews. Interviews were recorded on digital media and transferred to a computer. Files were then uploaded to a secure web-based password-protected server. Interviews were transcribed verbatim and the resulting text documents uploaded to the same web server. The interview protocol was followed as provided with additional questions generated by the interviewer to follow up on conversational points. In the first year of participation, fall telephone interviews were relatively brief and focused on teachers’ progress with the initial technology set-up and progress in implementation. The protocol used for fall telephone interviews is located in Appendix C. In subsequent years, fall telephone interviews were more comprehensive, probing aspects of pedagogy and technology implementation in greater detail.

Spring telephone interviews were conducted in May and June of 2007. All interviews were carried out by the same researcher who had conducted the fall interviews, thereby promoting the sense of relationship that participants have with project personnel. Files were recorded, uploaded, and transcribed as described for the fall telephone interviews. The interview protocol was followed as written (see Appendix D), with additional questions generated by the interviewer as needed to follow up on interesting aspects of the conversation. The same spring telephone interview protocol
was used for both years, since this instrument was also designed to collect specific data about the implementation of the technology.

*Classroom observations.* Based on their responses to the autumn telephone interviews, participants who had implemented the technology were selected for classroom observations. Members of the research team contacted the teachers to arrange for a convenient time to conduct the observation. Prior to the classroom observation, teachers were given a form to complete in which they described the classroom and available technology as well as the goals of the lesson to be observed.

Classroom episodes were videotaped with a digital recorder. In some cases, an external microphone was positioned near the front of the classroom to capture teacher comments as well as student interactions. The researcher operated the video camera so that teacher movement and behaviors were captured while students without parental consent did not appear on tape. After the observation, the video tape was converted to a digital file and uploaded to the secure web-based server. Verbatim transcripts were made by listening to the audio portion of the recording and viewing the video for clarification of complicated interactions.

*Teacher post-observation interviews.* While visiting each participant, researchers conducted a post-observation interview according to a prepared protocol, found in Appendix E. Additional questions that emerged as the researcher observed the lesson were added as appropriate. Post-observation interviews were audio-recorded on digital media, and uploaded and transcribed as indicated above.

*Student focus groups.* Prior to the observation or on the first day of the observation, students were given a short Likert-type scale questionnaire (see Appendix F)
about their experiences with the Navigator™. The results of these were tallied by the
observer and used in to guide the interviewer in conducting the student focus group
protocol (see Appendix G). Focus groups were conducted during lunch times or other
free periods so that students did not miss academic time. Participation in student focus
groups was solely on a volunteer basis. Students were given name placards and asked to
indicate a name for the researcher to use in the interview to preserve students’ anonymity.
Focus group sessions were audio-recorded, and uploaded and transcribed as described
above.

Summary

The CCMS project gathered a wide variety of both qualitative and quantitative
data from teachers and their students. Instruments and interview protocols were designed
to elicit information regarding aspects of the connected classroom experience from both
teacher and student perspectives. Data from the first year of the science portion of this
research study are used in the dissertation as described in Chapter 4.
Chapter 4 : Research Methodology – Data Analysis

The data included in the dissertation study presented here are taken from the first year of the physical science pilot study of the Connected Classroom Technology in Promoting Mathematics and Science Achievement (CCMS) project. Chapter 3 of this dissertation describes the research project as well as the qualitative and quantitative measures used to collect data. All of the data used in this dissertation study were collected as part of the CCMS project, as described in Chapter 3. All relevant available data for each participant were analyzed in completing this dissertation study. In Chapter 4, the particular methods used to address the guiding question are outlined. The first section of this chapter presents demographic data for the participants chosen for the analyses conducted in this dissertation as well as some background regarding their technology prowess, educational background, and context of their teaching environment. In the second section, methods and coding systems for analyzing and interpreting the previously collected data for the dissertation study are described. Finally, additional considerations regarding the conduct of this dissertation study are presented, including issues of trustworthiness.
Data Collection

Participants

Ten participants from the CCMS project attended the Summer Institute for training in the use of the TI-Navigator™ system and related pedagogy in the summer of 2006. Of these, four were able to implement the technology in their classrooms and returned to teach science for a second year with the TI-Navigator™ system. These four participants form the basis of the dissertation study presented here.

Participants were asked to complete demographic questionnaires (located in Appendix A). As shown in Table 4.1, the teachers whose practice will be examined more closely in this dissertation study have a widely varying background in education. Teachers’ academic preparation ranges from one participant with an undergraduate degree in education and no graduate degree, to a content specialist who earned a doctorate in physics before entering the teaching profession. Two of the teachers are relatively new to the profession, having taught for two or three years before participating in the CCMS project, while the other two are seasoned master teachers with more than 15 years experience and graduate degrees in curriculum and instruction. One teacher had a student teacher in her classroom during the latter portion of the school year studied in this dissertation; the student teacher was not present for any interviews and demographic data is not available for him, although he was present and co-taught during several of the observed lessons.
Participants completed self-ratings regarding their general computer skills and use of educational technology at the outset of their participation in the CCMS research study (instrument located in Appendix B). The participants in this study, with the exception of Mrs. Roberts, are quite comfortable with technology in general (see Table 4.2). This participant also demonstrated markedly less confidence in her content background compared to the other participants, which makes sense given that she had majored in elementary education.

Specific items on this survey reported in Table 4.2 inquired about teachers’ frequency of attending general pedagogical as well as technology-specific professional development. For the most part, the teachers in this study have not had a great deal of ongoing professional development. A noticeable exception is Mrs. Woods. In inspecting the raw responses from the survey, one finds that this participant has attended many graphing calculator workshops and is a frequent workshop presenter as well.

<table>
<thead>
<tr>
<th></th>
<th>State</th>
<th>Geographic Designation</th>
<th>UG Degree</th>
<th>Grad Degree</th>
<th>Years teaching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mr. Daniels</td>
<td>PA</td>
<td>Urban Fringe of a Large City</td>
<td>Physics, 1982</td>
<td>M.S, Ph.D. Physics, 1987</td>
<td>2</td>
</tr>
<tr>
<td>Mrs. Roberts</td>
<td>OH</td>
<td>Mid-size city</td>
<td>Elementary Ed, 1987</td>
<td>Curriculum and Instruction, 1995</td>
<td>16</td>
</tr>
<tr>
<td>Mrs. Black</td>
<td>OH</td>
<td>Urban Fringe of a Large City</td>
<td>Secondary Education, 2003</td>
<td>In Progress</td>
<td>3</td>
</tr>
<tr>
<td>Mrs. Woods</td>
<td>VA</td>
<td>Urban Fringe of a Large City</td>
<td>Geophysics, 1984</td>
<td>Curriculum and Instruction, 2004</td>
<td>21</td>
</tr>
</tbody>
</table>
Table 4.2: Technology and Content Background of Participants

<table>
<thead>
<tr>
<th></th>
<th>General Computer Skills*</th>
<th>Educational Technology*</th>
<th>Content Background*</th>
<th>Professional Development**</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mr. Daniels</td>
<td>3.23</td>
<td>3.00</td>
<td>4.61</td>
<td>2.20</td>
</tr>
<tr>
<td>Mrs. Roberts</td>
<td>2.55</td>
<td>2.55</td>
<td>2.43</td>
<td>2.20</td>
</tr>
<tr>
<td>Mrs. Black</td>
<td>3.77</td>
<td>3.76</td>
<td>4.52</td>
<td>2.20</td>
</tr>
<tr>
<td>Mrs. Woods</td>
<td>3.23</td>
<td>3.10</td>
<td>3.83</td>
<td>3.50</td>
</tr>
<tr>
<td>Science***</td>
<td>3.54</td>
<td>3.28</td>
<td>3.66</td>
<td>2.88</td>
</tr>
</tbody>
</table>

Higher scores indicate greater comfort level; scale 1-5  
**Higher scores indicate more frequent participation; scale 1-5  
***Average score of science teachers (n=19) comprising initial sample of CCMS participants

The settings in which these teachers work differ as well (see Table 4.3). One participant teaches classes with high numbers of special education students in a middle school, while another teaches physics to older high school students. The schools are all relatively large, but the socioeconomic background of students differs markedly.

Table 4.3: Demographic Characteristics of School Settings.

<table>
<thead>
<tr>
<th>School Enrollment</th>
<th>% Low SES</th>
<th>Racial Composition</th>
<th>Average Pretest Score</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mr. Daniels</td>
<td>1795</td>
<td>90</td>
<td>90% White 1% Hispanic</td>
<td>66.5%</td>
</tr>
<tr>
<td>Mrs. Roberts</td>
<td>820</td>
<td>50</td>
<td>57% White 1% Hispanic</td>
<td>41.2%</td>
</tr>
<tr>
<td>Mrs. Black</td>
<td>896</td>
<td>5</td>
<td>98% White 1% Black</td>
<td>53.8%</td>
</tr>
<tr>
<td>Mrs. Woods</td>
<td>1227</td>
<td>8</td>
<td>89% White 3% Asian</td>
<td>81.7%</td>
</tr>
</tbody>
</table>
Data Analysis

The previous section described the four participants whose classrooms and interviews make up the data corpus for this dissertation. In the following section, the methodology used to analyze these data will be presented. Each of three main types of collected data (teacher and student interviews, student surveys, and classroom observations) is described in terms of the assertions that the data support. Coding schemes are presented for each claim and are located in the Appendices of this dissertation.

Analysis of Interview Data

Telephone interviews were conducted in autumn of 2006 as well as spring of 2007. Two-day classroom observation visits occurred in the winter or spring of 2007, accompanied by post-observation interviews of the teacher as well as student focus groups. Interviews were conducted as described in Chapter 3. In this section, the coding schemes used to analyze a variety of interviews are described. The complete coding book is located in Appendix H. Coding of all interviews was performed using the software program NVivo, version 7. Interview data were used to support claims related to instructional tasks, teacher knowledge of student learning, and teacher instructional decision-making.

*What kinds of instructional tasks are implemented?* Interview transcripts were scrutinized for statements related to the use of connected classroom technology for instructional tasks. Formative assessment requires some kind of instructional task to be given to learners. Generally, teachers pose questions to find out how well students perform the task. The use of connected classroom technology involves the possibility of
new types of instructional tasks, but the nature of these tasks is not well characterized. The intervention in this research study specifically declined to impose any expectation of specific technology use on participants; rather, this study aims to uncover how teachers naturally integrate technology into their existing practices.

The first inspection of four interviews (student focus group, post-observation interview, fall telephone interview, and spring telephone interview) was used to identify all segments related to instructional tasks or activities. Instructional tasks and activities include any kind of classroom activity that a teacher implemented in order to teach a lesson or to assess for student understanding. By contrast, discussions of school-wide celebrations and classroom management procedures were not included as instructional tasks. A second analysis of the coded interviews sorted coded segments into one of four separate categories.

1. Non-examples: This category was used to identify segments of interviews that had been identified in the first analysis but upon further examination did not meet the criteria for task.

2. Hypothetical examples: In interviews, teachers occasionally spoke about their intended plans for the use of connected classroom technology. These were not included in further coding because it was unclear whether the plans were implemented.

3. General: The general category was used to categorize instructional tasks that were discussed in vague terms, such as “give quizzes” or “find out what students were thinking.” It also included vague statements about
lessons, such as “I used [connected classroom technology] when we were doing the waves unit.”

4. Specific: Specific tasks included clear descriptions of individual lessons or instructional strategies. To be included in this category, a description needed to mention the specific content or topic and/or a more detailed discussion of an instructional strategy than was included in the general category.

After categorization of tasks, those segments of interviews that fell into the general or specific categories were tabulated and divided further into types of instructional tasks. These codes included assessment (diagnostic, summative, and district/state related), reviewing concepts, monitoring for comprehension, and data collection/display. Further analysis resulted in collapsing and expanding these categories when discrete assignment of tasks to a single category was not possible. The final five categories used are: classroom management and behavior, summative assessment, preparation for state/district assessment, data collection and display, and review/check for understanding.

Do teachers know more about student learning? One of the primary purposes of formative assessment is for teachers to know more about student learning, in order to change instruction accordingly. Four different interviews for each teacher were inspected with the purpose of identifying comments related to the teacher’s knowledge of student understanding. Interview segments related to student grading, overall class progress, specific student misconceptions, and individual student learning were classified.
What instructional decisions do teachers make? As part of the formative assessment process, the information gathered by teachers should inform their instructional decision making process. Interviews for all participants were examined for evidence to support types of decisions made by teachers. Mentions of instructional decisions included no change in instruction as well as general review and reteach comments. Additional categories of instructional decisions included designing specific activities to address student misconceptions as well as individualizing student remediation.

Analysis of Student Observation Surveys

Student observation surveys were administered to students in advance of the two-day classroom observations. Sixteen Likert-scale items in matched pairs queried students about their perceptions regarding instruction and student-teacher interactions in a connected classroom. One item, “The TI-Navigator helps the teacher tell if I understand a concept,” was deemed useful for the analysis presented in this dissertation. Responses to all survey questions were entered into a spreadsheet, using values of 1-5 to correspond to ratings of strongly disagree, disagree, neutral, agree, and strongly agree, respectively.

Data for all four classrooms included in this dissertation were transferred to SPSS version 14.0. A t-test was used to compare the mean response to a neutral value of 3. An average score greater than three reflects overall agreement with the item.

Analysis of Classroom Observation Transcripts

The main source of data for describing assessment episodes comes from videotape and/or audiotape of classroom observations. Between 150 and 250 minutes of taped classroom lessons, comprising at least three discrete class periods per teacher, was
analyzed for each of the four participants from the 2006-2007 academic year. All available classroom videotape and/or audio recordings, approximately 800 minutes total, were included in the analyses described below. Recordings were transcribed verbatim and verified for accuracy while watching the video recordings.

*How are assessment episodes defined?* A major goal of this portion of this dissertation is to define and categorize the types of classroom discourse that occur with respect to assessment. As defined here, an assessment episode consists of a protracted series of prompts, questions, and responses that all relate to a single topic or assigned task. An assessment episode consists of the posing of a single question and the extended conversation that occurs based on that question. Typical class periods consist of one or more instructional tasks. In discussing instructional tasks, the teacher and students will engage in one or more assessment episodes. Assessment episodes can be of varying lengths, and no set number of assessment episodes must comprise the discussion related to a particular instructional task.

Assessment episodes were coded directly from transcripts of classroom observations. The beginning of an assessment episode is marked by a question, directed either at the teacher or at students. The ending of an assessment episode occurs when an answer to that initiating question has been received and responded to, or when an extended chain of questioning has reached a conclusion. The difference between a new assessment episode beginning and an existing assessment episode continuing for another round of questioning lies in the relationship between the subsequent questions and the initiating question. If the subsequent questions relate to the same line of questioning, probing for additional information or clarifying responses, they are part of the same
assessment episode. Conversely, if subsequent questions demonstrate a shift in the questioning process and no longer relate to the initial question, they are part of a new assessment episode. A question that can stand alone, not in relation to a previous question, begins a new assessment episode. Further, assessment episodes are frequently separated by the specific skill or content objective being assessed in the episode. Questioning patterns that are directed at materials management and other classroom procedures are not considered assessment episodes; similarly, discussions of non-academic topics are not part of assessment episodes.

Assessment episodes were not identified for small-group conversations. Teacher-student interactions occurring in small-group settings were not distinct enough in audio and video recordings to be able to capture with reasonable fidelity. Therefore, assessment episodes were only determined for whole-group discourse. The total number of assessment episodes was tabulated for each lesson. Since class periods were of different lengths for each teacher, and teachers spent different proportions of class time in whole-group or small-group instruction in the different lessons, this tabulation was then adjusted for the length of whole-group instruction, according to the procedure described below. Classroom observation recordings were viewed using computer software that included a recording time counter. Whole-group instruction was counted as the time from the beginning of the class period until students were dismissed to small group laboratory work, and then resumed when the class was called back to attention. Such transitions were indicated by teacher statements as well as by student movement in the room. Frequency of assessment episodes was then determined by dividing the number of assessment episodes in each class period by the total number of minutes of whole-group
instruction. In a few cases, students worked on small-group activities for the entire class period, and no assessment episodes were recorded.

Within each assessment episode, teacher-student discourse occurred in predictable *patterns of questions, responses, and follow-up or feedback*, referred to here as QRF cycles. Although an initiating question could arise either from the teacher or from a student, the pattern of assessment episodes typically occurred as a cycle of a *question* posed by the teacher, a *response* from the student, and some type of *follow-up* action by the teacher. Assessment episodes that ended after a single cycle of questioning, responding, and following up were termed single-iteration QRF cycles. Multiple-iteration QRF cycles occurred when the follow-up action was the posing of further questions and the assessment episode continued. The *degree* of the assessment episode is then determined by the number of iterations of QRF sequences; a single-iteration QRF cycle has a degree of one, while a pattern consisting of three iterations of QRF sequences has a degree of three.

*What patterns of questioning, responding, and following-up occur?*

Within assessment episodes, the specific types of questions, responses, and follow-up were categorized in detail. The foundation for the initial coding of QRF cycles arose from the work of Chin (2006), who describes an analytic framework for describing student-teacher interactions in middle school science classrooms in Singapore. In this framework, Chin extends the traditional notion of an IRE/F (initiation, response, evaluation or follow-up) cycle to a more detailed examination of the purposes and types of utterances at each stage. Coding of individual questions, responses, and follow-ups proceeded using an analytic induction technique, with *a priori* codes based on Chin’s
(2006) work. Revision of codes by modifying, eliminating, and adding categories was performed by the primary researcher, with coding by a second coder as described in a subsequent section of this chapter, *Trustworthiness*.

The first stage of an assessment episode is the *initial question*. Every assessment episode begins with an initial question. Because initial questions set out tasks for students to complete, Bloom’s taxonomy was chosen for an initial coding scheme. In Bloom’s taxonomy, *Knowledge* questions are at the lowest cognitive level, followed by *Comprehension* and then *Application* questions. Higher order questions consist of *Analysis*, *Synthesis*, and *Evaluation*; these three types of question were combined into one coding category of *Higher Order Questions* for the study reported here. An additional type of question, the *Rhetorical* question, does not require a student response. Due to the resultant low cognitive demand, *Rhetorical* questions are grouped with *Knowledge* questions.

The second component of a QRF cycle consists of a response to the initial question. Response categories included *no response* during the allowed wait time, *off-topic* and *tangential responses*, *non-verbal responses*, *yes/no responses*, *words and phrases*, and *explanations, hypotheses or conjectures*. Some responses were inaudible from the video recordings; based on the follow-up comments from the teacher, these were classified to be words and phrases from the student, since a common follow-up from the teacher was to repeat the students’ responses.

In multiple-iteration assessment episodes, the teacher follow-up step is to ask an additional, related question. For multiple-iteration QRF cycles, these *subsequent questions* address student responses and probe student thinking in various ways.
Categories for subsequent questions include *implied* questions, *redirected* questions, *clarifications and repetitions*, and *requests for elaboration*. Codes for all phases of QRF cycles and their descriptions are included in the coding book in Appendix H.

Other Considerations

*Trustworthiness*

The dissertation work presented here is primarily of a qualitative nature. In qualitative work, certain standards of trustworthiness must be attained. Lincoln and Guba (1985) describe four aspects of trustworthiness as including how confident others may be in the credibility of a researcher’s findings, the applicability to other contexts, the consistency if the study were to be replicated, and the neutrality of the researcher(s). In this section, aspects of this dissertation study that address trustworthiness and credibility are described. With a small sample size that was purposively selected for successful implementation of technology in secondary science classrooms, no claims for transferability or generalization are being made. The primary researcher is unlikely to be entirely neutral, and potential sources of bias are discussed in the next section, “The Researcher as Instrument.”

Triangulation of methods, data sources, researchers, and theoretical frameworks can all be used to establish the credibility of a qualitative study (Lincoln & Guba, 1985). In this dissertation study, several data sources are employed. To characterize assessment episodes and student-teacher interactions, two days of classroom observation for multiple class periods have been analyzed; this reduces the possibility that one extreme situation is being used as the basis for the findings presented as part of this dissertation. Further,
student focus group interviews, teacher interviews, and student surveys are being used to understand the extent to which teachers know more about student learning from different perspectives and at different times of year. The use of multiple forms of data therefore increase the trustworthiness of the findings.

In addition, multiple coders have been employed in the analysis of the data. The primary researcher established *a priori* codes for the qualitative analyses to be conducted in this study. Two major portions of the coding were deemed to require external checks on the coding process, while other minor tasks were essentially clerical and did not utilize second coders. The first major analysis that used a second coder was the breakdown of classroom observation transcripts into individual assessment episodes. The second coder for this step was the primary researcher’s advisor and co-Principal Investigator for the research study from which the data are drawn. A complete transcript of one classroom observation was selected and given to the second coder along with the initial coding book containing definitions and descriptions of assessment episodes. Coding was performed using NVivo, version 7.0. The primary researcher prepared a comparison chart of the two sets of codes in order to examine the differences that arose. Complete overlap of coding was obtained for assessment episodes related to science content and science processes, although some discrepancy was noted for teacher and student questioning behaviors that involved classroom management. Coded segments of the transcript that contained classroom management or other types of “teacher talk” were discarded, and this eliminated the discrepancy. Further analysis of the science content and process assessment episodes compared where the two coders began and ended each assessment episode; differences in assignment of these points resulted in redefinition of assessment
episodes in the coding book. The primary researcher completed the remaining breakdown of transcripts into assessment episodes.

The second aspect of classroom observation coding that used a second coder was the categorization of questions, responses, and follow-ups present in assessment episodes. The additional coder for QRF cycles has recently completed a doctorate in science education and has over 15 years of classroom and administrative experience in K-12 schools. A total of 24 assessment episodes (8.2%) were coded using the initial coding book developed by the primary researcher. Of 126 discrete questions, responses, and follow-ups, the initial coding process resulted in disagreement for 29 lines, or 23% of the coding. Discrepancies were discussed and resolved where possible. After resolution, 14 lines (11.1%) of coding were still discrepant.

Modifications to the initial coding book were made, resulting in the final coding book reported previously and located at the end of this dissertation in Appendix F. In several cases, definitions were expanded and/or reworded. Several categories were collapsed due to a lack of reliable distinction between the coding categories. One area of discrepancy that was unresolved between the two coders involved the assignment of a code to one particular pattern of questioning. This occurred where the teacher-assigned task was for the class to review student responses presented on the board. The intent, as indicated by the teacher and according to other assessment episodes within the same classroom activity, was to have students evaluate each other’s responses. The initial questions in this circumstance were therefore coded as Higher Order by the first coder. However, in ten of the assessment episodes related to this instructional strategy, the initial question was phrased as a yes/no question. It was the strong opinion of one coder
that a question to which an accepted answer is a “yes” or a “no” response cannot be a higher-order question, and coded such questions as Comprehension. For the purpose of the analysis presented in this dissertation, the discrepancy and argument is well noted, and the final decision of the primary researcher is to maintain coding based on the intent of the question in this particular circumstance. The reasoning for this decision includes the notion that if a student has evaluated a response and found it to be correct, he may simply respond with a “yes.” The primary researcher completed the remaining coding of QRF cycles within assessment episodes.

The Researcher as Instrument

As a researcher, I bring certain experiences and biases to this work. As a high school science teacher for nearly 9 years, I am sharply attuned to many aspects of classroom practice. Although this can be a benefit in terms of my observational skills, it also means that I might focus only on aspects of classroom practice that resound with my own experience. I am more likely to interpret classroom activities in light of my own experiences.

While in the classroom, I endeavored to implement technology in a number of different ways. Again, these experiences provide me with a better idea of the kinds of challenges that teachers experience when implementing technology, but they also predispose me to the possibility of projecting my own experiences onto the participants in this study. For a variety of reasons, primarily related to the infrastructure of my school and curricular demands, I was unsuccessful in implementing innovative technologies the way I wanted to.
I was a student observer/participant in the weeklong Summer Institute that all four of these participants attended. Although I did not spend as much time with them as the other participants did, I was able to get to know these individuals on a personal basis before becoming involved with the CCMS project as a research assistant. I have visited all of their classrooms and interviewed all of these teachers. Moreover, I have listened to their interviews, watched their videos and read the transcripts for data collected by other researchers. I see these participants at the yearly professional development follow-up sessions held each February. While this ongoing relationship can be beneficial, as a researcher I will need to remain vigilant that I am analyzing their data with an appropriate level of detachment, using the insights gained from my immersion in their classroom practices but respecting their individuality and rights.

Summary

This chapter described the specific analysis of data presented in this dissertation study. The four participants whose classrooms and experiences comprise the data for this dissertation were introduced along with descriptive statistics regarding their education backgrounds, technology expertise, and school contexts. The different qualitative and quantitative measures administered as part of the study were described, along with specific items of interest for quantitative measures. Assessment episodes were defined, and the codes used to characterize the questions, responses, and follow-ups within these assessment episodes were presented.
Chapter 5: Results

The use of interactive classroom technologies, as discussed in Chapter 2, can be an effective component of a supportive learning environment. The role of connected classroom technology in uncovering student learning needs makes it a strong potential match with formative assessment practices. The link between interactive classroom technologies and formative assessment remains to be clarified.

Previous studies indicate that teachers believe they know more about student learning through the use of connected classroom technology (Irving et al., 2009; Pape et al., 2008). This suggests that a classroom implementing this type of technology provides a richer environment for teaching and learning. These classrooms are therefore likely to be a fruitful site for examining formative assessment practice. Current definitions of formative assessment center on models of formative assessment practice where a teacher poses an instructional task, gathers assessment data in order to know more about student learning, and uses that information to make appropriate changes to instruction in order to increase student achievement. The guiding question of this dissertation is: How can rich formative assessment practice be described in detail? This study will provide a close analysis of connected classroom environments in order to shed light on a new model for formative assessment by addressing these focus questions:

1. What instructional tasks provide rich assessment information?
2. How can classroom-based assessment episodes be analyzed in detail while retaining the context of the related science instruction?

3. What characteristics of assessment episodes support rich formative assessment practice?
   a. What types of questions are used at the beginning of assessment episodes?
   b. What types of student responses are elicited?
   c. What strategies are used in subsequent questions to probe student understanding?

4. How do teachers gather data about student learning?

5. How do teachers use formative assessment data to make instructional decisions?

In this chapter, data collected from teacher interviews, classroom observations, and student surveys will be presented in the context of defining a more detailed model of formative assessment practice. This model takes into account many of the facets of the formative assessment cycle presented previously and arises from the analysis of data from connected classrooms. Each of five main assertions consists of an aspect of formative assessment analyzed in greater detail, with examples from teacher practice to serve as benchmarks on a continuum of formative assessment practice. These assertions are combined to generate a more complete model of formative assessment in whole-class discourse in connected classrooms. Each assertion includes a presentation of how connected classroom technology contributes to rich formative assessment practice.

Depending on their level of expertise, teachers would be expected to be at different stages of formative assessment practice. Even the most experienced teacher
would not be expected to be always performing at the “exemplary” end of the formative assessment continua presented in this chapter. Teachers make their instructional decisions taking into account a wide variety of inputs; in some cases, it may serve other instructional purposes better to not practice rich formative assessment. However, for teachers who wish to improve their FA practice, this model serves as a signpost for self-monitoring of practice as well as a guide for professional development. This model also provides an organizing instrument for outside observers, such as curriculum designers, university supervisors, teacher mentors and the like, to examine closely formative assessment practice.

Defining Rich Formative Assessment Practice

This chapter describes the results of a close analysis of teacher and student interviews, student surveys, and classroom observations in order to examine aspects of formative assessment practice. Five assertions are presented, each one contributing understanding to a different component of formative assessment. Each assertion concludes with a portion of a complete model for rich formative assessment practice, the whole of which is presented at the conclusion of this chapter. The overall model of formative assessment that emerges from this analysis answers the guiding question: How can rich formative assessment practice be described in detail?

Assertion One: Rich Instructional Tasks are Implemented

The first step in formative assessment cycles is for the teacher to set the instructional task and communicate the criteria by which it will be assessed. In this section, the varieties of instructional tasks reported by participants in their first year of
implementing connected classroom technology are summarized. Then, findings from interviews with teachers and students are used to identify rich instructional tasks that support teachers learning about student thinking. At the end of this section, categories of instructional tasks are summarized and presented along a continuum of formative assessment practice. The data presented under this assertion answer the focus question:

What instructional tasks provide rich assessment information?

What science topics and tasks can be implemented through connected classroom technology? Teachers in this study reported using connected classroom technology to support instruction in all four of the major science disciplines (chemistry/physical science, earth and space science, life science, and physics/physical science). Observed lessons covered mainly physics-based topics (waves, electricity, and mechanics) with one set of observations of a chemistry-based physical science lesson (acids and bases). Table 5.1 displays the variety of lesson topics that teachers reported implementing with CCT.

Lessons were implemented with a range of student ages and backgrounds, as outlined in Chapter 4. An important aspect of the intervention in this research study is that teachers are not required to implement specific lessons or activities. Although the summer workshop introduced myriad potential ways to use this connected classroom technology in science classrooms, any actual uses of the Navigator system in these classrooms reflect the individual teacher’s comfort level and personalization for his or her teaching style and curriculum.
Teachers expressed differing purposes for their implementation of connected classroom technology, as seen in Table 5.2. Some of these purposes, such as reviewing and checking for understanding as well as data collection and aggregation, are more closely aligned with strong formative assessment practice. Others, such as classroom management, serve little formative assessment purpose but remain essential components of the classroom environment and are a reality of teaching practice.

Table 5.2. Purposes of Instructional Tasks Implemented Through CCT

<table>
<thead>
<tr>
<th>Participant</th>
<th>R/U*</th>
<th>DC*</th>
<th>PA*</th>
<th>SA*</th>
<th>CM*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mr. Daniels</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mrs. Roberts</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Mrs. Black</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
<td>x</td>
</tr>
<tr>
<td>Mrs. Woods</td>
<td>x</td>
<td>x</td>
<td></td>
<td>x</td>
<td>x</td>
</tr>
</tbody>
</table>

Using connected classroom technology for review and checking for understanding can occur in a variety of ways. Several teachers identified the use of technology to administer brief checks for understanding, frequently ungraded, at the beginnings of class periods. Referred to as “do-nows” or “bell-ringers,” these assessments usually reviewed material from a previous class session and were used by the teacher to make sure that students had reached a certain level of understanding before proceeding with the new lesson. Mrs. Woods describes how she incorporated a “do-now” before a particular lesson using an application that allows her to send questions electronically to students’ handheld calculators:

For instance, today … they logged right on, and the room was quiet as they started doing their Learn Check. Then they got to some later ones and they were discussing it, but the talk was about the topic … The questions that I use for the Learn Check, just today, were questions that last year I might have used out of their book. I just modified them for Learn Check. (Mrs. Woods: fall telephone interview)

Although the items came from a teacher ancillary source, she did modify and enter the particular questions that were most relevant to the topic.

Similarly, Mrs. Roberts uses connected classroom technology to have students solve formula-based problems about motion. She describes how the discussion about the class’s responses on the Navigator screen helped students to make sense of their calculations:

Especially when we were doing velocity problems and there was math. And my kids are not strong math students. And they would put answers in. And then when we were looking at them on the slide show, kids would go, oh, they multiplied instead of divided. They were able to look at people’s responses and figure out what they did wrong. (Mrs. Roberts: spring telephone interview)
The interactive nature of this system allows the teacher and students to uncover aspects of student learning that might otherwise have gone unnoticed.

The preceding examples of instructional tasks supported by connected classroom technology primarily involve checking for understanding during lectures and class discussions. However, a central feature of science instruction is participation in laboratory exercises and the collection and manipulation of related data. The particular connected classroom system implemented in classrooms participating in the CCMS research study has features that facilitate the use of data in a variety of ways. Teachers in this study describe the following aspects of laboratory data analysis:

- Collecting student data
- Pooling or aggregating class data sets
- Displaying individual or pooled data sets as graphs
- Having students generate best fit lines
- Sending aggregated data to students’ calculators for advanced analysis

Often in science classes, teachers will have small groups of students collect different parts of a data set in order to save time and to allow student choice in a guided inquiry context. This then requires students to pool their data so that each student has a complete data set with which to carry out further analysis. Mrs. Woods appreciates the ease with which individual data sets can be shared among all students, telling the interviewer that “I used it in labs … so that people would have more than just the data they have gathered so they could share data” (Mrs. Woods: spring telephone interview). Similarly, Mrs. Black speaks about the time saved by aggregating data electronically:
I used Navigator to collect their pH values, and we calculated the class average. So instead of having to have 12 kids come up to my chalkboard, they just sent it to me, I could collect it real fast, and I had the second person in each group with the second calculator figure out the average. It took a lot less time, I thought. (Mrs. Black: post-observation interview)

Collecting, aggregating, and manipulating laboratory data is an essential component of science instruction that is facilitated using connected classroom technology.

A different type of data regarding students, that of student performance on statewide high-stakes achievement tests, has become a guiding factor in curriculum and instruction in many school districts. In some school districts, additional internally-developed or third-party assessments are used to monitor student progress and prepare teachers and students for the accountability tests. As one type of formative assessment practice, two participants mention using the Navigator system to review for their state-and/or district-wide science achievement testing programs.

An example of this use occurs when Mrs. Roberts “keep[s] a database of OAT [Ohio 8th Grade Achievement Test] questions that my district has published for the teachers as review questions … we do a thing called a daily start-up, so what I do is … send it as a Quick Poll to begin class” (Mrs. Roberts: spring telephone interview). This teacher also relates her students’ request to use the Navigator system when conducting a review for quarterly district-level achievement tests, as she describes to an interviewer.

My 8th graders have a common assessment that we take … four times a year. We have a big review day coming up on Monday. [The students] said, can we review with the Navigator and do the Learn Checks, because we pay more attention to it. Rather than doing a study guide or anything. And I said, sure. (Mrs. Roberts: fall telephone interview)
Similarly, Mrs. Black “did a pre-OGT [Ohio Graduation Test] … I had them put their answers in Learn Check and then I put all four class results onto one Class Analysis, so that we could get a breakdown of each question for all four classes combined” (Mrs. Black: fall telephone interview). In this case, the other teachers in the department also chose to administer the same test using the Navigator system and used the results to identify student strengths and weaknesses before the high-stakes test was given.

Depending on the ways in which teachers and departments use the information collected through practice assessments and official high-stakes tests, these instructional tasks can be considered as a component of formative assessment. In these two cases, the use of high-stakes testing for diagnosis of weak areas of performance and subsequent curricular changes may be considered a type of formative assessment.

In addition to using connected classroom technology to administer practice assessments for high-stakes testing, several teachers incorporated CCT into their regular classroom practice to facilitate grading. In some cases, the limited number of calculators in the classroom presented a challenge to administering a quiz. Mr. Daniels explains how he circumvented this difficulty by providing an alternate task for students not taking the quiz at that time:

Interviewer: Did you use these for actual grades for students?
Mr. Daniels: I did I think maybe two times, I think I did.
Interviewer: And you managed that by splitting the class in half?
Mr. Daniels: Yeah, yeah, half the class would take the quiz while the other half worked on an activity and then I would switch them up. (Mr. Daniels: spring telephone interview)

Student response to the use of connected classroom technology to administer summative assessments was mixed. Many students reported that they liked seeing their scores right
away. One of Mrs. Black’s students remarked that “[the teacher] would be able to tell us our grades just right after the test instead of having to wait like a couple of days for her to grade it with her pen” (student focus group; Mrs. Black’s classes).

However, in a few cases, students expressed a preference for taking tests and quizzes in the traditional fashion. One of Mrs. Black’s students tells the interviewer that:

Female Student: I liked it better when we just do it on paper instead of with a calculator.
Interviewer: Why?
Female Student: I do not know. It just seems like it is right in front of you instead of having to look and messing up or something. I just like it better when you can write it down. (student focus group; Mrs. Black’s classes)

This student evidently feels more comfortable testing in traditional ways, and her discomfort seems to stem from the physical manipulations on the handheld keypad and network connections. To her, the extra steps required to locate and choose a response outweigh the potential benefits of responding electronically. When probed, her classmates informed the interviewer that there were not usually technological issues with sending responses through the networked system. In fact, one young man in this class opines that “[responding on the calculator] is more prone to have mistakes as far as answers goes but if you are careful enough and care enough then you are going to be okay” (student focus group; Mrs. Black’s classes). He feels that, despite some infrequent calculator malfunctions, the user interface is straightforward and the communication errors do not misrepresent his responses. This particular teacher has students complete a written test, circling their responses on a paper version, as they respond electronically, which serves as a backup source of student grades if necessary.
A final set of instructional tasks can be categorized as “classroom management and lesson pacing.” A major use of the connected classroom technology within this category is related to monitoring students for off-task behavior and for completion of procedures. Students are able to install games on the handheld calculator devices used with this particular connected classroom technology. One teacher explained how she used a feature of CCT to generate a virtual snapshot of students’ calculator screens, described as a *Screen Capture*:

Some of them would go to playing games, and I wasn't aware of that, that they could do that, and that kind of threw me off. I would screen capture, and I would see games -- I would see stacking bricks. And so then I told them, I can [screen capture] at any time and see what you are doing. And I did notice, now that I’ll do that, they are less likely to go play the games. (Mrs. Roberts: post-observation interview)

Some less-frequent mentions of classroom management-related uses include eliciting responses from more students and identifying which students had not yet responded, as well as using the system as an attention-grabbing device.

Participants in this study used connected classroom technology in many different ways. These uses included techniques that facilitate classroom management and grading, which are critical components of strong pedagogy but do not provide much insight into student understanding. Other uses of technology included providing instructional tasks to students that challenged them to respond to a variety of prompts: manipulating experimental data, answering questions, solving problems, and identifying locations on images. Additional examples of uses of connected classroom technology are found in the following sections.
Teachers implement student-centered discussions. Tasks that allow the teacher to learn the most about student learning are those that encourage open student discussion and sharing of ideas. In science classrooms, this often occurs as teachers engage the class in analyzing collected laboratory data. Student discussions are particularly fruitful when data sets are displayed for all students in the form of graphs and equations of best-fit lines, as occurs with the use of the connected classroom technology implemented in this research study.

Mrs. Woods relates to an interviewer how she used connected classroom technology to encourage a rich discussion about the data collected in a laboratory exercise.

They collected data, just with some little toy trucks and stop watches … They sent it up, and we looked at it, and then I sent it all back, and they did a manual fit line. Then they sent their equations in for me. So, the first time they sent in points, you know, ordered pairs, and then the second time they sent in equations. We looked at their equations and talked about how they were similar and how they were different and figured out what the different variables meant. What did the Y intercept mean? What or how did it correspond to what they had just done and what did the slope mean? How did that correspond to what they had just done? (Mrs. Woods: fall telephone interview)

Here, students are conducting experiments related to motion, sharing their results, and discussing the application of their results to the physics concepts they have been learning. This student-centered instructional task provides many opportunities for the teacher to investigate aspects of student learning, either formally or informally.

Other examples of beginning class with planned assessments were aimed at preparing students for upcoming laboratory work and reviewing the results of experimentation. In this interview, Mrs. Woods shares her thoughts about the way she designed the lesson that was observed by the research team that day. She is describing a
lab exercise in which students use long coiled wire toys to model features of transverse waves.

I had a little plan, a little outline that I wrote out on a sheet of paper, because I had never done this lesson before … I knew I wanted them to review the parts of a wave, distinguish between frequency and amplitude, and wave length, and be able to choose which one were larger and smaller, and so on. I wanted them to think about it in terms of their lab. So, I wanted them to go from their lab, back to this picture. I intentionally asked questions like, “What happens when you move your hand faster?,” as opposed to “Which wave has a higher frequency?,” so they could try to relate it and then step up to the frequency. (Mrs. Woods: post-observation interview)

Her questions are not planned to elicit vocabulary terms and memorized facts; rather, Mrs. Woods has planned assessment items to help students make sense of their laboratory experience.

One participant in this study specifically mentions that the use of connected classroom technology to collect, aggregate, and display laboratory data allowed her to identify student difficulties. She contrasts this with the traditional method of individual student or small group data collection and graphing. The teacher tells the interviewer that:

From a teaching standpoint, you have kids do these labs and they only ever see their data or their group’s data and the one line they put on the graph. Or you have them come up to the overhead and add lines … [I] aggregate all of their data when they send it to me, and then I can put it into a graph and have all of their plots. And then we can pick out the data that is not valid. (Mrs. Roberts: fall telephone interview)

As an example of invalid data that became apparent to her, this teacher also relates a time when the aggregated data plots show an anomaly. Although this teacher usually chooses to keep student responses anonymous, she was able to identify the student and engage in the following conversation:
So I looked at the girl, and I said, do you know what you did? Look at this. And she said, oh, I put the temperature in before the time. And I said, yes, and what’s wrong with that? And she said, the x information goes first. And I did it backwards … But from that she learned what she was doing wrong. I would have never caught that if we were just having a classroom discussion. (Mrs. Roberts: fall telephone interview)

Although the use of networked handheld devices to collect data may appear to be limited to a data-management tool, episodes like this one demonstrate the potential of interactive technology to facilitate richer assessment and instructional practices.

*Teachers implement teacher-centered lesson delivery and verification tasks.* In many classrooms, the primary instructional task for students is to listen to a lecture and take notes. In a teacher-centered situation, fewer opportunities for probing student thinking, and therefore formative assessment, exist. Certain aspects of connected classroom technology allow the teacher to embed spontaneous questions in class discussions or lectures. Mrs. Roberts makes the distinction between planned and unplanned uses of questioning clear in her comment:

Sometimes I type in a question that I will have ahead of time and then I will just send that out to the students after my instruction. Sometimes I will just think of a question, so I will ask them, without typing it in and let them respond to me. (Mrs. Roberts: spring telephone interview)

Note the emphasis here on the teacher’s instruction and the questions that she chooses to administer to students. Similarly, Mrs. Roberts explains that she regularly has students complete “a daily start-up, so what I do is I keep a bank of those [Ohio Achievement Test] questions. And I will just pull one and send it as a Quick Poll to begin class” (Mrs. Roberts: spring telephone interview). Although a teacher can uncover student misunderstandings by using verification questions embedded in the lecture, this practice
is less informative than when students generate the questions and guide the discussion themselves.

Review and checking understanding with spontaneous unplanned questions allows the teacher to briefly pause instruction in order to get a response from all or most of the students. Mr. Daniels explains that “in the middle of a lesson [I] throw up a Quick Poll question about what they're doing and have them answer it” (Mr. Daniels: spring telephone interview). Mrs. Roberts describes a time when she gave a Learn Check to her seventh grade class.

They were doing organisms and individuals and populations, and they were working with a lot of different organisms in the classroom. So I did a Learn Check on that vocabulary, and I found that many kids, when asked about a population, they actually would write things like "a rock.” This was after we had practiced this vocabulary, and we had used the organisms in the classroom,… it was a misconception, it really was. (Mrs. Roberts: post-observation interview)

Although the students had been covering this material for several days, they still had misunderstandings, which the teacher may not have otherwise realized.

Mrs. Woods relates a specific instance where she used this technique to find out whether her students knew as much as she thought they knew:

Mrs. Woods: We were logged in already, and I was talking about something, and I did the bad habit of saying, “Okay, have you guys got it?” I look around and I saw these heads nodding and I said, “Great, I’m going to send you a Quick Poll.” I typed in a question and I got almost no correct answers … I said, “That’s okay. Just looking at your heads, I would have assumed that you got it, but now I know you don’t, and we can talk about it some more.”

Interviewer: How did they respond to that?
Mrs. Woods: They just laughed. They thought it was okay. So, we talked about it more, and then I sent another Quick Poll and most of them got it right. (Mrs. Woods: fall telephone interview)
In this episode, the teacher relates a time when she was providing direct instruction and felt that she could continue to go on based on how she interpreted students’ body language. After querying the class to verify their understanding of a concept, Mrs. Woods realized that they did not understand as well as she thought they did. Although this is not a student-centered example, it does provide the teacher with a limited amount of formative assessment information.

An area where many students struggle in science classes is in performing calculations. Several examples of instructional task in this first year of CCT implementation for these four teachers involved assessing student skills with manipulating formulas. Mr. Daniels explains his use of CCT to check his students’ proficiency with performing calculations when they “were re-looking at what the efficiency was and how to calculate the different amounts of energies from electrical and heat energies. The students were able to do calculations with the equations that they were given” (Mr. Daniels: post-observation interview). In this episode, he has sent questions to students in electronic files that are transmitted to their handheld devices; students solve the problems and return their responses through connected classroom technology. The teacher can then see how each student has performed the assigned task.

An innovative use of this particular connected classroom technology is the ability to import and have students interact with images, such as scientific diagrams. Mrs. Black describes several uses of these:

I did a brief chemistry lesson, where I was using the periodic table, and when I talked about the valence electrons and metals and non-metals, I just had them move to whatever they think it is. (Mrs. Black: post-observation interview)
We were talking about heat transfers from ice to water to steam, and I had a graph pulled up on the screen. And I asked, where is the ice melting into the water, and they would take their points down to the first leveled off part of the graph. Then I'd ask where water turns into steam, and they would move their points up to where steam is at the top of the graph. (Mrs. Black: fall telephone interview)

In both of these examples, Mrs. Black has set up her system so that students can use the cursor keys on their individual handheld calculators to maneuver an electronic marker on the display screen at the front of the classroom. Although the specific questions that the teacher asks about such an image may be written out in advance or generated on-the-spot, the use of a graphic background requires advance planning for the teacher to locate the image and prepare to import it into the presentation software. Once the image is imported, however, the teacher has the flexibility to ask whatever questions he or she finds most appropriate.

*Teachers emphasize behavior management and grading.* Instructional tasks and other teacher actions are administered for a variety of reasons, including for behavior management and grading of students. Not all of these purposes are suited to rich formative assessment practice.

For example, teachers may choose to administer quizzes using connected classroom technology but not share the results with their students. One of Mrs. Woods’ students explains this practice to the interviewer when he says “We have done a few quizzes on there, like smaller quizzes. She just takes it in, and we don’t get to see the answers on the board. She tells us what they are” (student focus group; Mrs. Woods’ classes). Without discussion of student responses or some other evidence of the teacher using the information to guide instruction, this cannot be considered rich formative assessment practice.
All of the participants in this study explicitly mention using connected classroom technology to administer and score summative assessments, in the form of short quizzes and tests. Teachers appreciated the ability to rapidly score quizzes and go over the results immediately. Mrs. Black tells the interviewer:

I get their grades instantly, so usually before they leave class that day I've had the answers up on the board and we've gone over it. So they get that instant feedback, they know what they did wrong or right, right away. (Mrs. Black: post-observation interview)

Mrs. Black’s emphasis on using connected classroom technology appears to focus on its utility in scoring student work rather than its potential to reveal deeper student understanding.

Connected classroom technology was also used for behavior management purposes, and the students quickly realized the potential of the networked system to reveal whether they were focused on their work. When a group of students was asked about off-task behavior, one student responded, “she can keep track of who is goofing off and who is not learning and then she can use it as a more effective method of disciplining us” (student focus group; Mrs. Black’s classes). The student has equated the notions of learning and discipline.

Another major use of connected classroom technology for classroom management purposes entailed the monitoring of student progress in completing procedures. Teachers used the technology to identify which students were finished with a given task or who needed more time to work. They also used connected classroom technology to ensure that students were using calculator applications properly. One participant tells the interviewer that she “would screen capture what they were typing in, and partly it was to
make sure they were typing their information into the right lists, to make sure they were under lists actually” (Mrs. Black: fall telephone interview).

By contrast, one of the participants remarks that using the technology to achieve this monitoring step did not always fit well with her instructional style:

I used Screen Capture just to see how people were progressing say in a Learn Check, see where they were to gauge when I could go on. But it took a while to refresh the screen, to get everyone's to load, and so once people were finishing I found it easier really just to walk around the room. (Mrs. Woods: spring telephone interview)

It may be that this experienced teacher has a strong enough sense of her classroom that the technology does not increase her ability to monitor student progress. Without this added benefit, the mechanics of using the connected classroom system actually impedes her instruction. Mrs. Woods had a classroom set of TI-83 Plus calculators, with less memory than the more advanced TI-84 Plus, and a slower laptop computer, adding to the inconvenience of using connected classroom technology for progress monitoring.

Although the two areas of task administration (summative assessments, classroom management) described above are important components of teaching and learning, they do not exemplify tasks that lead to rich formative assessment practice. The administration of summative assessments tends to focus on whether students have given correct or incorrect responses and how the teacher can use the gathered data to assign a grade to each student. Tasks related to monitoring inform the teacher about how well students understand classroom management procedures, rather than how well they understand the science content and processes being taught. They also help to keep students engaged in the lesson, in part by revealing when the student is off-task.
Few examples of classroom practices designed solely for behavior management and grading purposes were observed in this study. Presumably, this may be due in part to an “observer effect,” as teachers would be anxious to show off their best practices to a researcher visiting the classroom. However, information gathered about students during tasks designed to keep students engaged and on-task or designed to assign grades to students’ work is certainly useful to teachers in other aspects of their professional practice.

The three main types of instructional tasks discussed in this section are arranged in a continuum of practice in Figure 5.1. In this figure, the left-most benchmark indicates a minimal emphasis on formative assessment practice, while the right-most category describes exemplary formative assessment practice.

<table>
<thead>
<tr>
<th>Richness of Practice</th>
<th>Emerging Practice</th>
<th>Exemplary Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>How Instructional Tasks Are Used (Assertion 1)</td>
<td>Behavior management, Grading</td>
<td>Teacher-centered; content delivery</td>
</tr>
</tbody>
</table>

Figure 5.1. Use of instructional tasks

Assertion Two: Questioning Is Deep and Probing.

In this assertion, the notion of an assessment episode is introduced and defined. Summary statistics on the average duration and frequency of assessment episodes will be presented, along with a discussion of how cycles of questioning can occur within assessment episodes. Evidence presented in this section answers the focus question: How
can classroom-based assessment episodes be analyzed in detail while retaining the context of the related science instruction?

What is an assessment episode? Whole-group interaction is an effective setting in which teachers both instruct and assess regarding science concepts. Assessment Episodes (AEs) are defined here as the set of questions and answers that occur throughout a lesson. An assessment episode begins whenever the teacher or a student initiates a question on a new topic or subtopic. The assessment episode continues through recurring cycles of Questioning, Responding, and Following up, referred to here as “QRF cycles,” until a new line of questioning is initiated. Assessment episodes in the observed lessons of the four participants included in this dissertation were categorized and analyzed in greater detail to provide insight into the teachers’ formative assessment practices. Table 5.3 depicts the number of assessment episodes analyzed for each teacher as well as the average frequency of assessment episodes during whole-group instruction. Details on the classification of assessment episodes and determination of frequency are provided in Chapter 4 of this dissertation.

One might infer from these data that richer formative assessment occurs in classrooms where more assessment episodes occur. After all, a teacher cannot assess student progress without engaging in some type of data collection; an increased number of assessment episodes may indicate a greater amount of data regarding student understanding being collected. However, this dissertation will not attempt to support this claim for a variety of reasons. The data in Table 5.3 are presented merely for background information regarding the amount of whole-group instructional discourse that occurred in the observed lessons. The limited number of lessons observed for each teacher is not
likely to provide an adequate depiction of usual teacher practice. Furthermore, as subsequent sections of this analysis will demonstrate, the content of each of these assessment episodes varies greatly.

Table 5.3. Number and Duration of Assessment Episodes

<table>
<thead>
<tr>
<th>Participant</th>
<th>Total Class Time Analyzed (minutes)</th>
<th>Whole-Group Instructional Time (minutes)</th>
<th># Assessment Episodes</th>
<th>Average # AE per minute of whole-group instruction*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mr. Daniels</td>
<td>234</td>
<td>109</td>
<td>31</td>
<td>.15</td>
</tr>
<tr>
<td>Mrs. Roberts</td>
<td>182</td>
<td>108</td>
<td>123</td>
<td>.86</td>
</tr>
<tr>
<td>Mrs. Black</td>
<td>159</td>
<td>134</td>
<td>100</td>
<td>.82</td>
</tr>
<tr>
<td>Mrs. Woods</td>
<td>257</td>
<td>125</td>
<td>40</td>
<td>.32</td>
</tr>
<tr>
<td>Total</td>
<td>832</td>
<td>476</td>
<td>294</td>
<td></td>
</tr>
</tbody>
</table>

*calculated independently for each observed lesson and averaged for each teacher

What are single-iteration assessment episodes? At a minimum, an assessment episode consists of a single QRF cycle, where the teacher (or student) poses a question, student(s) respond, and the teacher follows up with some acknowledgement or feedback. The follow-up may entail the decision to ask a subsequent related question, thereby continuing the assessment episode, or to switch to a new instructional activity, closing the assessment episode. Depending on the outcome of this decision, an assessment episode
may consist of a single QRF cycle; alternatively, linked QRF cycles may be iterated some number of times.

As shown in Table 5.4, the majority of assessment episodes, both overall and for individual teachers, consist of single QRF cycles (58.2% overall). For example, in this episode from Mr. Daniels, the teacher asks a single question, a student responds with an incorrect statement, and the teacher provides corrective feedback to the class. This ends the assessment episode.

Teacher: (Q) How many paths for the electricity to flow does a series circuit have?  
Student: (R) More than one.  
Teacher: (F) A series circuit is a circuit that has only one path for electrons to flow. Okay? (Mr. Daniels; Per 1 Day 1)

Another example of a single-iteration QRF cycle occurs in this episode from Mrs. Roberts’ class. Here, the teacher is engaging students in a discussion of a hypothetical car trip during a lesson on motion and position-time graphs. This time, the teacher provides a question prompt, a student responds with a correct value, and the teacher affirms this response by repeating and elaborating on the student response.

Teacher: (Q) I can take that entire distance I went, maybe five miles and then five miles back home, which would be a total of what?  
Student: (R) Ten.  
Teacher: (F) Ten miles. (Mrs. Roberts; Per 1 Day 1)

Although commonly used, the effectiveness of a single-iteration QRF cycle is limited by its briefness. In a single-iteration QRF cycle, only one student is providing a response to the question prompt. Furthermore, in a single-iteration QRF cycle, the initial student response is accepted regardless of how well it addresses the intent of the teacher’s question. The student is “let off the hook” when the teacher stops the QRF cycle after
one iteration. Frequently, the teacher provides the desired but missing information from the student by elaborating on the given response. In the example above, by saying “ten miles” Mrs. Roberts provides the unit of measurement that was left out of the student response.

Table 5.4. Iterations of Assessment Episode by Teacher

<table>
<thead>
<tr>
<th>Participant</th>
<th>Number of Assessment Episodes</th>
<th>Number of Single-Iteration AEs (%)</th>
<th>Number of AEs with 2 or 3 Iterations (%)</th>
<th>Number of AEs with 4 or more iterations (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mr. Daniels</td>
<td>31</td>
<td>16 (51.6%)</td>
<td>12 (38.7%)</td>
<td>3 (9.7%)</td>
</tr>
<tr>
<td>Mrs. Roberts</td>
<td>123</td>
<td>71 (57.7%)</td>
<td>39 (31.7%)</td>
<td>13 (10.6%)</td>
</tr>
<tr>
<td>Mrs. Black</td>
<td>100</td>
<td>64 (64.0%)</td>
<td>32 (32.0%)</td>
<td>4 (4.0%)</td>
</tr>
<tr>
<td>Mrs. Woods</td>
<td>40</td>
<td>20 (50%)</td>
<td>17 (42.5%)</td>
<td>3 (7.5%)</td>
</tr>
<tr>
<td>Total</td>
<td>294</td>
<td>171 (58.2%)</td>
<td>100 (34.0%)</td>
<td>23 (7.8%)</td>
</tr>
</tbody>
</table>

*What is a multiple-iteration assessment episode?* With an increased number of iterations of QRF cycles, the teacher does not accept the first response given but rather requires students to provide additional information in some manner. For example, the teacher may prompt the class to elaborate on a prior response, as seen in the following excerpt. Here, Mrs. Black is asking students to give the charge of a particular family of ions. Ions can have either a positive or a negative charge, depending on their parent
atom’s electron configurations, and this is a critical distinction that is not provided by the student. Rather than accept “one” as a response or state the desired answer, this teacher prompts the class to elaborate on the answer. She then restates both parts of the answer to clarify the desired response. Contextually, students appear to understand that “what kind of one” refers to the positive or negative nature of the charge in question.

Teacher: (Q₁) So everything in the first column of the periodic table has what kind of charge?
Student: (R₁) One.
Teacher: (F₁) One; (Q₂) what kind of one?
Student: (R₂) Positive.
Teacher: (F₂) Positive one, good. (Mrs. Black; Per 3 Day 1)

Multiple iterations of the QRF cycle might also be used to have students clarify their thinking. In the next excerpt, Mrs. Roberts is asking students to predict how the line on a position-time graph will change when the motion of an object changes. The initial response provides a correct statement, but lacks a distinction that the teacher brings out by asking the student directly to compare this leg of a journey to a previously discussed stage.

Teacher: (Q₁) Then I get in my car and go to the store; what’s going to happen to the line?
Student: (R₁) It’s going to go up.
Teacher: (F₁) It’s going to start getting steeper again. (Q₂) Will it go as steep as it did the first leg of the trip?
Student: (R₂) No.
Teacher: (F₂) No. (Mrs. Roberts; Per 3 Day 1)

In addition to the more extensive probing demonstrated in two- and three-iteration QRF cycles, a small number of assessment episodes consist of four or more iterations. These include episodes where multiple students are queried and/or content raised in earlier portions of the assessment episode is investigated further. Assessment episodes
that contain more QRF cycles provide a means by which the teacher can delve into
student understanding, searching for nuances of understanding that are not apparent when
fewer-iteration QRF cycles are employed. In this example, Mrs. Roberts is asking
students to interpret why the position-time graph is structured differently when the
reference point is at a different relative location from a previous class example.

Teacher: (Q₁) Why is the reference point up here?
Students: (R₁) He went the opposite way. He started at the top.
Teacher: (F₁; implied) (Q₂) He started at the top?
Student: (R₂) I mean he started at [inaudible] instead of at home.
Teacher: (F₂) He didn’t start here at home; (Q₃) he started…
Student: (R₃) [inaudible]
Teacher: (F₃) A distance from home, (Q₄) right?
Student: (R₄) That’s what I meant.
Teacher: (F₄) So remember, on the Y is distance, so this is showing where you
start; it’s your starting point, ok? (Mrs. Roberts; Per 3 Day 1)

Here, several students have provided initial responses (R₁) that trouble the teacher, who
selects one response to focus on. She implies the request for students to clarify their
responses, which takes several attempts (Q₂, Q₃) before she provides appropriate wording
to use (F₃, Q₄).

Employing an increased number of QRF cycle iterations does not, in itself,
provide richer formative assessment data to a teacher. The additional cycles may reflect
a different instructional purpose, such as to elicit responses from a wider range of
students. One such case occurs when a teacher engages in a nine-iteration QRF cycle in
order to call on many students in the class to each provide one of the five possible single-
word responses to a question prompt. Here, Mr. Daniels is engaging in a review-game
activity where students must catch a soft toy in order to answer a question for
competition points. Some student utterances are related to the game rather than the content of the initial question.

Teacher: (Q₁) Stephanie, your question is Number 8, “What flows in a wire?”
Number 8: What flows in a wire?
Student: (R₁) Electricity? I do not know.
Teacher: (F₁) Okay. Electricity would be correct so I cannot say it is wrong. (Q₂)
   What flows in a wire, Curt?
Student: (R₂) Current.
Teacher: (F₂) Current, very good! (Q₃) Ed, what flows in a wire?
Student: (R₃) Electrons.
Teacher: (F₃) Electrons. Very good! There is one more I can think of. (Q₄; implied) Megan.
Student: (R₄) Amps.
Teacher: (F₄) Amps is current. Very good! (Q₅; implied) Kate.
Students: (R₅) Catch the ball, Kate! Volts?
Teacher: (F₅) All right. Kate said “Voltage.” Kate said voltage. (Q₆) Ray, would we say that is correct?
Student: (R₆) No!
Teacher: (F₆) Kate said “Voltage flows.” Voltage exists as energy; that is different points of a circuit. (Q₇) Finally, Alexia. I just said something else that flows in a circuit. I just said the word.
Student: (R₇) Electricity.
Teacher: (F₇) Did anybody catch that? (Q₈) What else flows in a circuit? To review: current, amps, electricity, joules.
Student: (R₈) Electrons.
Teacher: (F₈) Electrons. Julie, voltage does not. All right pretty good! Okay. (Q₉) Review. What flows in a circuit?
Student: (R₉) Electrons, electricity, watts.
Teacher: (F₉) Electrons, electricity, joules, amps, current, watts. (Mr. Daniels; Per. 1 Day 1)

Although more students are directly involved in this assessment episode, the repeated questioning cycles serve primarily to allow students to offer different low-level responses to a single low-level question prompt. Despite the assessment itself being at a relatively low level, the multiple-iteration cycle does allow the teacher to involve more students in the assessment.
The number of iterations of QRF cycles occurring within a single assessment episode is therefore an important component of rich formative assessment practice. Figure 5.2 depicts the role of QRF cycle iterations in supporting rich formative assessment practice. Assessment episodes consisting of a single QRF cycle allow the teacher to assess a single student on a single piece of knowledge, and are placed at the “emerging practice” end of the formative assessment continuum in this model. By contrast, greater numbers of QRF cycles within an assessment episode provide the teacher with the means to query multiple students about the same topic and/or probe an individual student’s understanding more deeply. Instructional practices that fit this pattern occur at the “exemplary practice” end of the model proposed in this dissertation.

<table>
<thead>
<tr>
<th>Richness of Practice</th>
<th>Emerging Practice</th>
<th>Exemplary Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Degrees of QRF Nesting (Assertion 2)</td>
<td>Single-degree QRF only</td>
<td>Some engagement in deeper probing; several degrees of QRF employed</td>
</tr>
</tbody>
</table>

Figure 5.2. Degree of assessment episodes

**Assertion Three: Classroom Interactions Elicit Higher-Order Thinking**

As mentioned above, the depth of QRF cycle iterations is only one component of a high-quality questioning strategy. An assessment episode that consists of the same low-level question repeated to many students does not provide the rich assessment information possible when a teacher engages in shorter assessment episodes that require students to think critically. A teacher engaging in single-iteration QRF cycles might be
asking students to use critical thinking skills in responding to a single higher-order prompt. Therefore, an important aspect of considering how questioning cycles contribute to formative assessment practices is to examine the quality of each part of the QRF cycle. In this section, three components to high-quality questioning cycles will be described. Initial questions, which begin an assessment episode; subsequent questions, which occur whenever an assessment episode contains a second or higher degree; and student responses are all considered as part of this assertion. The data presented here will address the focus question: What characteristics of assessment episodes support rich formative assessment practice? To accomplish this, the three components of assessment episodes described above are elaborated in turn, using the following focus questions:

a. What types of questions are used at the beginning of assessment episodes?

b. What strategies are used in subsequent questions to probe student understanding?

c. What types of student responses are elicited?

What types of questions are used at the beginning of assessment episodes? Each assessment episode, whether consisting of single- or multiple-iterations of the QRF cycle, begins with an initial question prompt. Higher-order questions, as defined here, include the analysis, synthesis, and evaluation levels from the original Bloom’s Taxonomy. More details regarding these classifications are described in Chapter 4 of this dissertation. An analysis of initial question prompts by participant is provided in Table 5.5.

For most of the participants, very few initial prompts consisted of higher-order questions. One teacher in particular, Mrs. Roberts, included a large fraction, over one-
fourth, of higher-order questions in her classroom practice (see Table 5.5). For instance, in the following episode she prompts the class to evaluate a student response to finding the speed of an object traveling 50 meters in 50 seconds.

Teacher: (Q) What’s wrong with this one?
Student: (R) The distance; it’s meters.
Teacher: (F) It’s the distance; good. (Mrs. Roberts; Per 1 Day 1)

By asking students to identify the incorrect aspects of a response, Mrs. Roberts is engaging in higher-order questioning. This in turn allows the teacher to uncover more of what students understand about a given concept; here, whether they realize that a response containing the correct value and an inappropriate distance unit is incorrect.

Since Mrs. Roberts’ lesson plan for the observed lessons was to have students respond and then discuss right and wrong answers, a high ratio of her initial question prompts would be expected to follow this evaluation pattern. Most higher-order questions across all participants were of the same format, suggesting that teachers have difficulty formulating higher-order questions for classroom discourse outside of the peer-critique instructional activity displayed in this example.

The largest number of initial questions in assessment episodes for all participants was in the application category (42.5% overall; see Table 5.5). Application questions specific to science content included performing calculations, interpreting a graph by describing the type of motion, and relating experimental procedures (hand motions) to an outcome (wave form of a Slinky™). One example of a participant assessing understanding with an application question is in the following excerpt.

Teacher: (Q) How do you calculate electrical energy?
Student: (R) You do Watts times seconds.
Teacher: (F) Watts times seconds. (Mr. Daniels; Per 1 Day 1)
Here, students have completed a lab where they collected data regarding the amount of energy required to operate household appliances.

### Table 5.5. Categorization of Initial Question Prompts by Participant

<table>
<thead>
<tr>
<th>Participant</th>
<th>Total Number</th>
<th>Rhetorical</th>
<th>Knowledge</th>
<th>Comprehension</th>
<th>Application</th>
<th>Higher Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mr. Daniels</td>
<td>31</td>
<td>0 (0.0%)</td>
<td>11 (35.5%)</td>
<td>8 (25.8%)</td>
<td>11 (35.5%)</td>
<td>1 (3.2%)</td>
</tr>
<tr>
<td>Mrs. Roberts</td>
<td>123</td>
<td>5 (4.1%)</td>
<td>13 (10.6%)</td>
<td>22 (18.7%)</td>
<td>50 (40.1%)</td>
<td>33 (26.8%)</td>
</tr>
<tr>
<td>Mrs. Black</td>
<td>100</td>
<td>4 (4.0%)</td>
<td>27 (27%)</td>
<td>29 (29.0%)</td>
<td>39 (39.0%)</td>
<td>1 (1.0%)</td>
</tr>
<tr>
<td>Mrs. Woods</td>
<td>40</td>
<td>0 (0.0%)</td>
<td>2 (5.0%)</td>
<td>7 (17.5%)</td>
<td>25 (62.5%)</td>
<td>6 (15.0%)</td>
</tr>
<tr>
<td>Total</td>
<td>294</td>
<td>9 (3.1%)</td>
<td>53 (18.0%)</td>
<td>66 (22.4%)</td>
<td>125 (42.5%)</td>
<td>41 (13.9%)</td>
</tr>
</tbody>
</table>

Another example of application questions occurs when the teacher asks students to predict an outcome, based on prior knowledge. In the following example, Mrs. Black is asking students to relate their knowledge of ion charges to the resulting charge when a substance becomes an ion when it accepts an additional atom.

Teacher: (Q₁) So if everything in the first column has a positive 1, we are going from water which is neutral, to what? We added that hydrogen, and what's its new charge going to be?
Students: (R₁) <no response>
Teacher: (F₁; implied) (Q₂) What was hydrogen’s charge?
Student: (R₂) <inaudible>
Teacher: (F₂) Positive 1, right? (Q₃) So if we add something that is zero, what's its new charge?
Student: (R₃) Positive one
Teacher: (F₃) Positive one. So now we are going to have this hydronium ion now, so instead of being called water we are changing it up and calling it hydronium. (Mrs. Black; Per 3 Day 1)
By asking an application question, the teacher is requiring students to think critically; they cannot just recite a piece of information they have memorized from a previous lesson. This higher level questioning allows the teacher to assess students at a higher level as she determines if they understand this rather abstract concept.

In a third example, Mrs. Woods is discovering whether students can apply their experiences in a laboratory exercise to the definitions and mechanics of simple transverse wave motion.

Teacher: (Q₁) All right, what happens to the wavelength if you shake it back and forth faster?
Student: (R₁) It gets shorter.
Teacher: (F₁) It gets shorter. (Q₂; implied)
Student: (R₂) It decreases.
Teacher: (F₂) Because you have more waves in the same amount of space, so the wavelength has to get smaller. (Mrs. Woods; Per 6 Day 1)

The first response from a student is repeated and thereby tacitly affirmed; the teacher waits for an additional response and elicits a clearer answer from another student.

“Shorter” is a term that might be confused with a description of amplitude rather than wavelength, which this question prompt addresses.

A sizeable fraction of initial prompts in this study was classified as comprehension-level questions (22.4%; see Table 5.5). Such questions require students to demonstrate an understanding of content that goes beyond recitation of facts, but does not show that they can use that information in a meaningful way. Mrs. Woods uses a comprehension question in the following assessment episode.

Teacher: (Q) But it is possible to make a low frequency, high amplitude wave; a high frequency, high amplitude wave; a low frequency, low amplitude wave, or what’s the other one?
Students: (R) High. Low.
Teacher: (F) [High] frequency, [low] amplitude wave. So you can have all those combinations because they’re all independent of each other. (Mrs. Woods; Per 7 Day 1)

Students need to understand that all these combinations of amplitude and frequency of waves can occur, as opposed to the dependence of wavelength and frequency. They are doing more than reciting a learned fact, such as providing the textbook definition of amplitude. However, they are not being asked to apply their knowledge.

The initial question in the remaining assessment episodes were low-level, either knowledge-based or rhetorical questions. Knowledge-based questions asked students to remember facts without providing any context, explanation, or application, and occurred in nearly 20% of assessment episodes (Table 5.5). Mrs. Black uses knowledge questions to make sure that students remember facts from a previous reading assignment. In this episode, students are asked to recall the range of pH values for acids and bases. Several responses are provided; the teacher asks students to repeat the response and a correct answer predominates.

Teacher: (Q₁) What part of the range deals with acids?
Students: (R₁) One through 5? Zero through six.
Teacher: (F₁; implied) (Q₂) What is it?
Student: (R₂) Zero through six.
Teacher: (F₂) Good. (Mrs. Black; Per 7 Day 1)

Low-level questioning provides the teacher with information regarding student recall of facts, which are essential as the components of more sophisticated thinking. However, if teachers incorporate an excessive number of knowledge-based questions, they fail to uncover how students are thinking about the content. Students’ revealed thinking is the factor that comprises strong evidence for formative assessment.
The remainder of low-level questions were categorized as rhetorical (3.1%; Table 5.5), where no response was required or expected from the students. In this excerpt from Mrs. Roberts, the teacher is reminding students of a previous lesson in which they experienced a related concept. The students do murmur their assent to having seen the graph being described, but provide no evidence that they understand the relationship between the experience and the current discussion.

Teacher: (Q) You guys even compared average speeds of different things. When you went to the computer lab and you picked a robot vs. a vacuum cleaner and you had those things, you ended up with two distance vs. time lines that looked like this, correct?
Students: (R) Um hum.
Teacher: (F) And all that shows you is how much distance over time each thing is moving. (Mrs. Roberts; Per 1 Day 1)

Rhetorical questions may serve a purpose in that they help the teacher to monitor student engagement. Furthermore, rhetorical questions assist in maintaining the pace of a lesson, but they provide little or no information regarding student comprehension.

The type of question that the teacher uses in posing an instructional task or assessment item to students is linked to the richness of formative assessment that can arise. Low-level questions that ask students to recall facts and terms inform the teacher only that the student has learned a response to a given question. Higher-order questions, on the other hand, allow the teacher to explore the extent of student understanding in detail. This in turn provides richer information on which the teacher bases any future instructional decisions. The relationship of initial questions to rich formative assessment practice is summarized in Figure 5.3 at the end of this section.

*What strategies are used in subsequent questions to probe student understanding?*

While a taxonomy of higher-order questioning can serve to classify the questions that
initiate QRF cycles, this categorization scheme does not prove useful in examining questions that occur later in multiple-iteration QRF cycles. A different taxonomy, then, must be used to understand more fully the nature of assessment episodes that consist of multiple iterations of the QRF cycle, which is nearly half of them (see Table 5.4). The question that initiates the QRF cycle is referred to as the “initial question.” In multiple-iteration QRF cycles, every iteration, by definition, includes a question of its own. In all of these, termed “subsequent questions,” the teacher presses students for additional information in some manner.

One type of subsequent question, the request for elaboration, is particularly useful in uncovering student understanding. Requests for elaboration comprised nearly 32% of all subsequent questions in analyzed assessment episodes (Table 5.6). In a request for elaboration, the teacher is acknowledging a student response but also attempting to get a slightly different response. The student has not given a piece of critical information, which the teacher hopes to elicit from the student.

An example of a request for elaboration occurs in the assessment episode presented below. The teacher has asked the class how to use an ammeter, which is a device for measuring electric current in a circuit. The teacher fields an off-topic response (R₁) and an incomplete response (R₂) before making the request for elaboration (Q₃).

Teacher: (Q₁) Number 12. How do you use an ammeter?
Male Student: (R₁) Oh, I know!
Teacher: (F₁) Nolan. No actually you do not. An ammeter is this. And there is no on or off switch. (Q₂) How do you use an ammeter? Ed?
Male Student: (R₂) Through a circuit.
Teacher: (F₂; implied) (Q₃) How do you test it?
Male Student: (R₃) In a series.
Teacher: (F₃) In a series. Very good! (Mr. Daniels; Per. 1 Day 1)
This is a request for elaboration because the elicited response, “through a circuit” is not incorrect, but does not provide a key piece of information that Mr. Daniels intends to elicit. An ammeter is indeed wired through a circuit, but the desired answer contrasts the use of an ammeter to the use of a voltmeter, which is wired in parallel to the circuit being tested.

Table 5.6. Categorization of Subsequent Questions

<table>
<thead>
<tr>
<th>Participant</th>
<th>Total Number</th>
<th>Implicit Leading</th>
<th>Repeat Verbatim or Ask Student to Repeat</th>
<th>Rephrase to Clarify</th>
<th>Reflect/Redirect to Student(s)</th>
<th>Request for Elaboration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mr. Daniels</td>
<td>36</td>
<td>10 (27.7%)</td>
<td>7 (19.4%)</td>
<td>5 (13.9%)</td>
<td>5 (13.9%)</td>
<td>9 (25.0%)</td>
</tr>
<tr>
<td>Mrs. Roberts</td>
<td>95</td>
<td>27 (28.4%)</td>
<td>6 (6.3%)</td>
<td>11 (11.6%)</td>
<td>9 (9.5%)</td>
<td>42 (44.2%)</td>
</tr>
<tr>
<td>Mrs. Black</td>
<td>57</td>
<td>17 (29.8%)</td>
<td>9 (15.8%)</td>
<td>11 (19.3%)</td>
<td>9 (15.8%)</td>
<td>11 (19.3%)</td>
</tr>
<tr>
<td>Mrs. Woods</td>
<td>28</td>
<td>7 (25.0%)</td>
<td>5 (17.9%)</td>
<td>8 (28.6%)</td>
<td>1 (3.6%)</td>
<td>7 (25.0%)</td>
</tr>
<tr>
<td>Total</td>
<td>216</td>
<td>61 (28.2%)</td>
<td>27 (12.5%)</td>
<td>35 (16.2%)</td>
<td>24 (11.1%)</td>
<td>69 (31.9%)</td>
</tr>
</tbody>
</table>

In another example of a request for elaboration, the teacher focuses on a piece of information given by the student. The teacher uses the response to probe more deeply into the response. An example of this occurs with Mrs. Roberts in reviewing how position-time graphs represent motion. The request for elaboration element in this assessment episode occurs at Q2.
Teacher: (Q₁) If we want to know the average speed we’re going to look for…
Student: (R₁) The slope.
Teacher: (F₁) The slope. (Q₂) Give me another word for that slope. We’re looking at the…
Student: (R₂) Line?
Teacher: (F₂; implied) (Q₃) How it’s what?
Student: (R₃) Inclined?
Teacher: (F₃) How steep it is, good (Mrs. Roberts; Per. 1 Day 2)

Here, Mrs. Roberts wants to be sure that students understand the link between speed and the steepness of a line depicted on a motion graph. The average speed is given by the slope, as indicated by R₁, but Mrs. Roberts draws additional information out of the students by prompting for further information in Q₂.

The response to a single question usually provides the teacher with information about one student’s understanding. Asking the entire class to answer a question is one way to elicit responses from all students, but the question and response do not necessarily represent critical thinking. A technique employed by teachers to bring more students into the conversation and to promote higher-order analytical and evaluative thinking is to reflect a student response back to another student or the class in general. As shown in Table 5.6, this occurred in 11.1% of the subsequent questions from multiple-iteration QRF cycles.

In the following assessment episode, the teacher has asked a question by sending it electronically to the students to answer on their handhelds. Before showing the responses, she asks students the question orally and receives several responses (R₁). She selects one student’s answer to reflect back to the class (Q₂).

Teacher: (Q₁) Positive motion; yesterday we talked about positive. Is positive motion moving away from the point of reference or the starting point?
Students: (R₁) False. True.
Teacher: (F₁) Lance says false; (Q₂) do you guys agree? What do you say?
Students: (R₂; responses on board)
Teacher: (F₂) 14 of you said true. Positive. You’re going away from the starting point. The line’s going to go up; it’s true. (Mrs. Roberts; Per. 3 Day 2)

By reflecting Lance’s response back to the class, the teacher is able to engage students and elicit their evaluations of a student response. Although this particular episode is unusual in that the individual student responses have already been given before the teacher reflects Lance’s oral response back to them, it illustrates the type of teacher move indicated by reflecting a response to individual students or to the class as a whole.

In another situation, Mrs. Roberts asks students to comment on where another student should initially stand in order to model the motion shown on a position-time graph. The student and teacher are demonstrating the use of a handheld motion sensor that will create a position-time graph of the student’s motion as he moves towards and away from the probe along a straight line.

Teacher: (Q₁) If this is where the trigger is, where do you think you should start? The X and Y intercept is where this actual motion detector is sitting, and the line is showing beginning here, where are you going to start?
Student: (R₁) You got to stop.
Teacher: (F₁; implied, non-verbal) Walk the line. (Q₂) Tell me about where you’re going to be when you start. Rashid’s trying to help you.
Student: (R₂) I’m not telling him.
Teacher: (F₂; implied) (Q₃) What do you guys think? You think he’s good?
Student: (R₃) Yup.
Teacher: (F₃) Start there? (Mrs. Roberts; Per 1 Day 2)

The student has made a decision about where to start his motion. Rather than confirm his choice or correct his position, the teacher asks the rest of the class to consider and evaluate his starting location. This serves to involve students, if vicariously, in the demonstration and have them test their understanding of reference points. To respond to a reflected prompt requires students to consider not only the initial question but also the
response under discussion. This causes reflecting responses to the class to be a higher level of interaction and questioning, and provides the teacher with more information about more students’ deep understanding.

In some cases, the teacher is able to tell from a response that students did not understand what the question was asking them to do. In the next iteration of the QRF cycle in such a circumstance would be termed a “rephrase for clarification” type of subsequent question. These make up 16.2% of the subsequent questions identified in Table 5.6. In the following assessment episode, Mrs. Black is asking students to determine whether a substance (bleach, in this case), is an acid or a base. Students are expected to base this judgment on the pH value they have just determined for bleach. The first student response (R₁) cannot be distinguished, but from the teacher’s next move it seems as though the inaudible response was not an expected response.

Teacher: (Q₁) What is bleach?
Student: (R₁) <inaudible>
Teacher: (F₁; implied) (Q₂) Is bleach an acid or a base?
Student: (R₂) <inaudible>
Teacher: (F₂) It is a base. (Mrs. Black; Per. 3 Day 1)

The initial prompt had been rather vague; “What is bleach,” for which there could be a number of potentially acceptable responses. In Q₂, Mrs. Black rephrases the question to narrow students’ possible responses to being either acid or base.

Rephrasing for clarification is important for formative assessment purposes because it allows the students to demonstrate their level of understanding. If students do not understand the questions asked of them, they are unlikely to be able to give a coherent response. Asking a question in a new form indicates that the teacher has used the student’s previous response to adjust instruction with additional guidance. The
teacher has learned that the student does not fully understand the question, so she is trying to elicit student knowledge in another way.

Other subsequent question types do not demonstrate that the teacher has used student responses to guide instruction. These include repeating the question verbatim, giving leading questions, and implied questions. Verbatim repeats of the initial question, 12.5% of all subsequent questions (Table 5.6), show that the teacher believes that students did not hear the initial question, students are purposely stalling for additional time to consider a response, or perhaps enough wait time has elapsed that the students may have forgotten the question. In this next assessment episode, Mr. Daniels repeats (Q₂) his initial question prompt after a student provides an incorrect response (R₁).

Teacher: (Q₁) What does an ammeter do?
Student: (R₁) It measures the wire.
Teacher: (F₁) No. It measures something. (Q₂) What does an ammeter do?
Student: (R₂) Measures volts? Amps?
Teacher: (F₂) Amps. Ammeter; think amps; current. (Mr. Daniels; Per. 1 Day 1)

Repeating the question verbatim in this case may assist students who are following along with a printed review sheet. However, it does not take advantage of a student response to probe into what the student is thinking, and therefore does not reflect strong formative assessment practice.

In classroom discourse situations, teachers may not always vocalize the repeated question, but imply that they wish to have the initial question answered. In assessment episodes, implied questions are recognized when a multiple-iteration QRF cycle occurs but no explicit question is vocalized between the feedback to one response and an additional student response. An example of implied questions is in the following
assessment episode. Here, Mrs. Black is asking students to recall from their reading assignment that the “p” in the abbreviation “pH” means.

Teacher: (Q₁) What does the “p” stand for?
Student: (R₁) Partly
Teacher: (F₁) Partly? That is close, good guess. (Q₂; implied)
Student: (R₂) Particles?
Teacher: (F₂) Particle, another good guess. Ok, it's actually power. we are detecting the power of hydrogen. (Mrs. Black; Per. 3 Day 1)

In this exchange, the teacher recognizes an incorrect response (R₁) and provides a gentle disconfirmation. She does not repeat the question explicitly (Q₂) but the students know that they should continue offering responses (R₂). Implied questions are also not strong examples of formative assessment, since the teacher is not adapting instruction or questioning techniques to meet student learning needs. Together with leading questions, where the student is expected to guess at the word that the teacher is looking for, implicit subsequent questions make up 28.2% of all subsequent questions.

In assessment episodes that extend past a single QRF cycle, the quality of questioning in subsequent cycles is related to the strength of the formative assessment possible. Subsequent questions that probe at student understanding, asking students to substantiate their claims, give teachers greater insight into how students are constructing their view of scientific phenomena. Questions that do not lend themselves to this type of mental exploration also do not provide the teacher with adequate information about student learning, thereby preventing the teacher from implementing high-quality formative assessment.

What types of student responses are elicited? A major goal in science education is to help students articulate their models of scientific phenomena. Student responses that
provide explanations or propose hypotheses inform the teacher about students’ ability to
describe scientific models, and therefore support rich formative assessment. As seen in
Table 5.7, responses including explanations and conjectures occurred in 12.6% of student
responses.

Table 5.7. Categorization of Responses in Assessment Episodes

<table>
<thead>
<tr>
<th>Participant</th>
<th>Total Number*</th>
<th>None Off-Topic</th>
<th>Tangential Nonverbal</th>
<th>Yes/No “I Don’t Know” Request to Repeat</th>
<th>Word or Phrase On Board</th>
<th>Conjecture Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mr. Daniels</td>
<td>78</td>
<td>8 (10.3%)</td>
<td>6 (7.7%)</td>
<td>4 (5.1%)</td>
<td>52 (66.6%)</td>
<td>8 (10.3%)</td>
</tr>
<tr>
<td>Mrs. Roberts</td>
<td>217</td>
<td>4 (1.8%)</td>
<td>3 (1.4%)</td>
<td>53 (24.4%)</td>
<td>112 (51.6%)</td>
<td>45 (20.7%)</td>
</tr>
<tr>
<td>Mrs. Black</td>
<td>166</td>
<td>7 (4.2%)</td>
<td>5 (3.0%)</td>
<td>17 (10.2%)</td>
<td>128 (77.1%)</td>
<td>9 (5.4%)</td>
</tr>
<tr>
<td>Mrs. Woods</td>
<td>71</td>
<td>3 (4.2%)</td>
<td>1 (1.4%)</td>
<td>8 (11.3%)</td>
<td>54 (76.1%)</td>
<td>5 (7.0%)</td>
</tr>
<tr>
<td>Total</td>
<td>532</td>
<td>22 (4.1%)</td>
<td>15 (2.8%)</td>
<td>82 (15.4%)</td>
<td>346 (65.0%)</td>
<td>67 (12.6%)</td>
</tr>
</tbody>
</table>

Note: total number of responses does not equal the number of QRF cycle
iterations because some student responses may be represented in more
than one category

An example of a student proposing an explanation occurs in the following
assessment episode taken from Mrs. Woods’ class. Students have been asked to identify
which of two waveforms displayed on a screen would result from a person shaking a
coiled metal toy spring with a greater amplitude.

Teacher: (Q1) Last time we did one where you shook your hand faster. This time
what would make – if you shook your hand farther away? Like one time
you shake your hand like this, and one time you shake your hand like this.
Which graph would represent that? And don’t yell it out.
Students: (R₁; selecting choices on screen)  
Teacher: (F₁) Ok, (Q₂) which graph represents when you shook your hand farther.  
Students: (R₂) The green one. The blue one.  
Teacher: (F₂; implied) (Q₃) Why the green one?  
Student: (R₃) It’s the blue one. It’s the blue one, because it goes higher.  
Teacher: (F₃) If you take the slinky and you shake your hand more your waves are going to become bigger. And they’re going to have a bigger amplitude because you’re giving that slinky more energy, so that’s going to give you the taller one. (Mrs. Woods; Per. 4 Day 2)

After students make their choices by moving an electronic marker with their networked handheld devices (R₁), the teacher asks students to respond to the repeated question (Q₂) verbally (R₂). When prompted for an explanation (Q₃), a student supports his response (R₃) with the evidence that the blue waveform has higher peaks. This exchange informs the teacher that the student has made the appropriate connection between the amplitude of shaking the metal coil and the amplitude of the waveform generated.

In addition to providing explanations, high-order student responses include proposing hypotheses and conjectures. In this assessment episode, Mrs. Roberts has asked (Q) her students to consider how a motion graph would display the return portion of a journey.

Teacher: (Q) Here’s what I want you to think about. What’s going to happen when I get in my car after the drugstore and I go back home?  
Students: (R) It’s going to get steeper and straight. It will get longer and steeper? Maybe?  
Teacher: (F) I want you to think about that. (Mrs. Roberts; Per 3 Day 1)

Here, the students (R) are proposing several conjectures for how the motion graph might display the return trip. Later in the class period (not shown here), Mrs. Roberts returns to this question and has students model the journey. By having students articulate their thoughts about how motion relates to the graphical depiction of motion, she learns what misconceptions they still have.

124
As opposed to the richer explanations and conjectures discussed above, students often respond to a teacher’s question with short phrases or single words. In this study, 65% of all student responses consist of a word or phrase (Table 5.7). Although words and short phrases do inform the teacher whether a student knows how to answer a particular question, in themselves they do not share nearly as much about student learning as longer explanations do.

Teacher: (Q₁) Everything in the first column of the periodic table has what kind of charge?
Student: (R₁) One.
Teacher: (F₁) One; (Q₂) what kind of one?
Student: (R₂) Positive.
Teacher: (F₂) Positive one. (Mrs. Black; Per. 7 Day 1)

In this exchange, it takes two rounds of questioning for the teacher to get to the desired response, and she still has to put the two parts (R₁, R₂) of the complete answer together for students. The student still has not vocalized the complete phrase signifying the charge of an element in the first column of the periodic table.

Later in the same class period, Mrs. Black asks students about their understanding of pH paper.

Teacher: (Q) Does anybody know how this (holding pH paper) works? What’s the scale; what’s the range?
Student: (R) Zero to 14.
Teacher: (F) Good; zero to 14. (Mrs. Black; Per. 7 Day 1)

A more complete phrase allows the teacher to know that the student understands both the starting and ending values of the pH range, as well as that a range requires both endpoints to be articulated.

Student responses that might otherwise be single words or short phrases can be, through the use of connected classroom technology, transformed to visual displays when
projected. Responses from the entire class are shared, rather than the comments of a few students. In the assessment episode below, Mrs. Black has asked students to use the arrow keys on their handheld devices to move an individual marker on the screen. She has displayed a color-coded pH scale, and student move their markers to indicate the pH value they believe a substance to have.

Teacher: (Q) All right, so it seems to me that most of you think borax is what, an acid or a base?
Students (chorus): (R) Acid.
Teacher: (F) An acid? Okay, so let’s see if you are right. (Mrs. Black; Per. 1 Day 1)

Mrs. Black can see at a glance what everyone’s response is. In this particular episode, she also chooses to ask students to vocalize what the majority of responses are.

Using the same software application, Mrs. Woods has students identify the wave with the shortest wavelength on an image containing several waveforms.

Teacher: (Q) For this one, go to the graph that has the shortest wavelength.
Students: (R; moving markers on screen)
Teacher: (F) It looks like most of you chose the red one. (Mrs. Woods; Per. 4 Day 2)

She provides students with sufficient time to move their markers to indicate their choice, and then summarizes the responses for the class.

Both of these episodes involved students making selections by moving a marker on the screen. Teachers using the type of connected classroom technology implemented in this study can also choose to have students input words and short phrases using the keypad on the handheld device. Here, Mrs. Roberts has used this feature to have students input the formula for calculating the distance traveled by an object.
Teacher: (Q₁) Nine of you said distance isXF minus XI; is that correct?
Student: (R₁) Yes.
Teacher: (F₁; implied) (Q₂) What’s XF, Lance?
Student: (R₂) I didn’t [inaudible].
Teacher: (F₂) That’s ok; (Q₃) what’s XI?
Student: (R₃) The final distance?
Teacher: (F₃) Good, and subtracting XI; (Q₄) what’s XI?
Student: (R₄) Initial.
Teacher: (F₄) Initial which means start; good. (Mrs. Roberts; Per 3 Day 1)

The assessment episode presented here occurs after students have independently submitted their responses and the software program has aggregated similar responses. The teacher is focusing on one response that was shared by nine students (Q₁). She uses this as a springboard for discussing whether the response was correct or not. The first student called upon to respond appears to be concerned that he had not been one of the students who submitted this particular response (R₁), but the teacher encourages him to answer her new question regardless (Q₃). These episodes demonstrate how connected classroom technology can promote class discussions that allow the teacher to uncover more about every student’s understanding.

Another group of student responses provides the teacher with limited information about student understanding. One such response is a yes/no response, or potentially a true/false response. These often follow very basic kinds of questions; higher-order questions generally cannot be answered by a simple choice of a yes or a no. In the following example, Mrs. Black has asked a leading question about the reactivity of atoms.

Teacher: (Q) Something that really, really wants to get rid of an electron, is it going to be reactive?
Student: (R) Yes.
Teacher: (F) Absolutely. And that’s why acids are so reactive. (Mrs. Black; Per. 7 Day 1)
The student provides a “yes” response (R), indicating to the teacher that this student agrees with her suggestion about reactivity. However, the student may have guessed, or the yes/no response may not have been necessary for the teacher to continue with her lecture. Other examples of student responses that provide limited information about student understanding include requests for the teacher to repeat the question, which indicate that the student may have missed the question or not understood it, and statements that the student does not know the answer. An example of the latter occurs in the assessment episode below. Here, Mrs. Roberts is asking a particular student to evaluate a student answer that is on the board. Initially, the student responds (R₁) that she does not know.

Teacher:  (Q₁) Jeanette, what was wrong with the last one?  
Student:  (R₁) I don’t know.  
Teacher:  (F₁) You don’t know.  (Q₂) It says ‘one sec.’  
Student:  (R₂) [inaudible]  
Teacher:  (F₂) Good; it should be one meter for every second; good, so you have to put the speed formula.  (Mrs. Roberts; Per 3 Day 1)

The student’s initial response does provide some formative information to the teacher; the teacher would certainly now know that the student doesn’t know the answer to this particular question. However, it does not inform the teacher about any potential confusions that the student might be having. With skilled prompting (Q₂), this teacher is able to elicit a better response from the student, as indicated by the teacher’s affirmation (F₂). As shown in Table 5.7, yes/no responses, indications of not knowing the answer, and asking the teacher to repeat the question together make up 15.4% of the student responses analyzed.
A final category of student response includes those that are off-topic or tangentially related to the topic under discussion. These are classified under the model presented here but do not contribute much to the formative assessment process, since they do not inform the teacher about student understanding. They may provide some information about student engagement and other classroom management issues that may influence instruction but do not directly indicate how much a student knows. Combined with no verbal answer, these responses make up 6.9% of all student responses analyzed in assessment episodes.

Considering levels of student responses is an important feature of a study of formative assessment practice. Although formative assessment is a teaching practice, the student responses comprise much of the data set that a teacher uses to make instructional decisions. If student responses are not of a quality that allows the teacher to make appropriate judgments regarding student learning, he or she cannot adjust his or her teaching accordingly.

The overall quality of questioning is linked to the richness of formative assessment practice. Teachers with more sophisticated questioning skills are able to encourage students to articulate their thoughts about scientific phenomena and concepts being discussed, rather than using questioning as a classroom management technique. Similarly, lower levels of student responses limit how much the teacher knows about student learning. Better evidence, leading to stronger formative assessment practice, is obtained when teachers ask deep, probing questions and receive thorough responses, as shown in Figure 5.3.
Figure 5.3. Quality of assessment episodes in rich formative assessment

**Assertion Four: Teachers Know What Their Students Understand**

As described above, formative assessment cycles include engaging learners in tasks that allow the teacher to collect data about the students’ learning in order to modify instruction accordingly. Recognition that the teacher knows more about student learning is therefore an essential component of formative assessment practice. In this section, responses from both the teacher and the student perspective will be used to describe ways in which teachers know more about student learning. Teacher awareness of student understanding addresses the focus question: *How do teachers gather data about student learning?* Teacher awareness of student learning is described from the student perspective as well as the teacher perspective.

*Do students believe that the teacher is aware of student understanding?* Table 5.8 shows the average response by students on a brief survey administered during or just before a classroom observation visit by the research team. One of the 16 items on this survey addresses the teacher’s awareness of student learning. Students in all four settings
agreed that the use of connected classroom technology assisted their teacher in knowing what students know.

Table 5.8. Student Survey Responses Regarding Teacher Knowledge of Learning

<table>
<thead>
<tr>
<th>Item</th>
<th>Mrs. Woods</th>
<th>Mr. Daniels</th>
<th>Mrs. Roberts</th>
<th>Mrs. Black</th>
</tr>
</thead>
<tbody>
<tr>
<td>The TI-Navigator helps the teacher tell if I understand a concept</td>
<td>3.60*</td>
<td>3.66**</td>
<td>4.03*</td>
<td>3.58*</td>
</tr>
</tbody>
</table>

* Mean score is different from a neutral value (3) on a 5-point Likert scale; p<0.01
** Value is significant at p=0.054

This question was followed up during the student focus group:

Interviewer: Are there other examples where you think that she has learned about what you are learning about physics by using Navigator?

Student: I mean, I assume there are some, but I really can’t think of what. Because lots of times, before she even like turns up an answer, before she knows what people respond with, she already goes over the part that everyone ended up missing … She’ll say that she is sure a lot of people missed this, and then she teaches us, and then she shows it, and a lot of people did miss it. (student focus group; Mrs. Woods’ classes)

The student who is speaking has noticed that the teacher does not rely on the gathered data to make her instructional decisions. The participant in question has taught physics for many years, and her decision-making process most likely reflects her years of accumulated mental data regarding common student difficulties rather than the data provided through the Navigator system.
Interestingly, the lack of a clear consensus from students regarding their teacher’s knowledge of their learning was apparent in the student focus group for Mrs. Black’s classes as well. When probed, Mrs. Black’s students responded:

Interviewer: The teacher knows just as much about my understanding without the TI Navigator as with it?
Female Student: Well, I do not really know, I mean, it is just hard to say. I guess yes, I agree with it. Yes.
Male Student 1: I actually disagree because I think with it the teachers can notice who is having trouble and who is not.
Male Student 2: She asks a question and you go to it and if you do not go to the right answer, know the right answer, she can click on it and figure out who it is and decide or tell whether you are having problems. And sometimes in a regular class, she is saying stuff and asks a question about [something], one person will give a right answer and someone else may not understand it but just hear the right answer and might not get it. (student focus group; Mrs. Black’s classes)

Some students feel that the teacher knows more about their learning using the technology while others disagree. Moreover, the focus from these students reflects an emphasis from the teacher on answering questions correctly and identifying incorrect responses, rather than an awareness of deep conceptual understanding.

Responses from participant interviews provide insight into the ways in which using a networked assessment system helps teachers know more about what their students know. In general, teachers felt that they knew more about their students through using this technology. Mr. Daniels mentions that “it just allowed me to have them respond to certain questions that I needed to know what they knew, there was a better way to get the information than just have [the students] respond orally” (Mr. Daniels: spring telephone interview). One way in which connected classroom technology allows teachers to know more about student learning is that all students have the opportunity, and in many cases
are required, to respond. A student in Mrs. Black’s class explains his perspective on how the use of connected classroom technology encourages all students to respond:

I have noticed that in this class is that when she asks a question maybe only like one or two kids, a.k.a. the smart kids, will answer and usually they would be the ones answering most of the questions most of the time. So when she does ask a question on the Smart Board [with the Navigator], everyone has to go to a question or to an answer. (student focus group; Mrs. Black’s classes)

The teacher knows what each student knows. Awareness of student understanding is critical for effective formative assessment. In the best case, the teacher would have evidence for what each student understands about a specific topic. Skilled use of connected classroom technology can facilitate the teacher’s uncovering each student’s level of understanding. Mrs. Roberts relates an instance that is indicative of her typical practice. She has sent students electronic questions to answer on their handheld devices and send back to her computer. The software allows her to identify specific student responses, and she uses this data to identify (“grab”) students who need some form of remediation.

But, we did a Learn Check … I didn’t share that, I didn’t put it up on the screen, but I was able to look on mine and see who was making those mistakes still. And I was able to grab them. So it helped me because I could pinpoint without embarrassing them at this point, who was still mixing those concepts up. I try to use that as much as possible. (Mrs. Roberts: fall telephone interview)

Mrs. Roberts’ awareness of specific student difficulties exemplifies formative assessment practice as she gathers evidence for each student’s learning and acts on that information to provide additional instruction.

The teacher knows about specific class-wide misconceptions. Often, during whole-class instruction, a teacher is more likely to get a sense of what students can and cannot do as a class. Group-wide formative assessment also benefits student learning as
the teacher modifies the group instruction to ensure that common difficulties are addressed. Mrs. Roberts tells the interviewer of her experiences with collecting student responses through a rapid polling technique and displaying student responses for the class to see and discuss.

Mrs. Roberts: I do a lot of open response. Because then we can look at all the different ideas that kids put up there and we can go through and discuss them. It’s good learning at the end of the discussion.

Interviewer: So what have you learned about your classes as a result of doing that?

Mrs. Roberts: I have taught for 18 years and I have been in 7th grade science for about 15 of the 18… and there are things that I have always been really sure that I think kids have understood completely. Now I see what they are thinking. And I am like, whoa, I am just amazed. (Mrs. Roberts: spring telephone interview)

Mrs. Roberts recognizes the importance of having a wide variety of responses from a large portion of the class in order to identify learning difficulties along the way. Despite her many years of experience and familiarity with teaching this content, this master teacher is finding out things about student learning that she had never realized before.

Another participant tells the interviewer of some specific content-related misconceptions held by the class that were uncovered through the use of questioning mediated through connected classroom technology.

Interviewer: What have you learned about your students from those questions?

Mrs. Woods: Well, one thing that I learned is that most of them really got what we talked about yesterday, which was very nice to see … I wouldn’t have expected as many of them to get it as did. Another thing that I have learned is that the whole concept of a medium is more confusing than I had thought it would be. Not only are they having a difficult time even deciding what to put down, they are discussing it amongst themselves, and they are really not convinced. Some of the kids, that usually get the right answer, think the wave goes through the air. It has been an interesting way to discuss it and rephrase what a medium is. What do we mean by “what the energy travels through”? (Mrs. Woods: post-observation interview)
In the lesson being described here, students had been sent electronic quizzes, answered questions on their handheld devices, and sent them to the teacher’s computer. During the ensuing class discussion, it became apparent that students were struggling with using waves to describe the concept of energy traveling through a medium. The teacher expresses surprise at this realization, but is not indicating in this excerpt how she might adjust her instruction in the coming days to address this conceptual difficulty shared by many students.

*The teacher has a general sense of class-wide understanding.* In some cases, the teacher may not be able to identify to such a level of detail what the misconception is. Often, a teacher has a general sense of whether students understand or not, rather than rich evidence for specific misconceptions. One participant relates her experience with using connected classroom technology to uncover student understanding in advance of quizzes and tests.

> I can get instant feedback, what they do and don't understand. So I guess it's allowed me to know what's going on in the middle of a class as opposed to finding it out later on a test or a quiz or because someone happened to answer a question.  (Mrs. Woods: spring telephone interview)

In this particular example, no evidence of a specific misconception is related, but the teacher is aware of the general level of understanding.

Similarly, Mr. Daniels explains how using connected classroom technology can help him as well as his students find out how the class responds to prompts.

> I think the biggest help with the Navigator was letting students see the typical questions and answers that everybody could see at the same time and respond to. And they get the feeling of being able to answer, and seeing how they answered, and me seeing how they answered. So that helped, I think.  (Mr. Daniels: spring telephone interview)
Again, no sense of specific misconceptions is being uncovered here. However, the teacher is aware that some students may not have answered in the desired manner. This is a step towards rich formative assessment practice.

_The teacher focuses on grades and summative assessments_. The reality of classroom settings is that teachers must pay attention to grades and summative assessments. The emphasis on grades in many classes and schools reinforces the importance of grades over understanding for students. Richer formative assessment practice most likely will not replace grading and summative assessment entirely. However, an excessive focus on grading and summative assessment may be indicative of weak formative assessment practice. A focus on grading and summative assessment may be apparent in interviews as teachers discuss learning in terms of grades and whether students provide correct answers. In this excerpt from a post-observation interview, Mrs. Black explains how she uses an application that allows her to send electronic quizzes to students, collect responses, and assign grades rapidly. The emphasis in this excerpt is on the convenience for the teacher of automatic grading, as well as the students, rather than the learning.

Taking a quiz for [the students] on Navigator is very simple…. I get their grades instantly, so usually before they leave class that day I've had the answers up on the board and we've gone over it. So they get that instant feedback, they know what they did wrong or right, right away. .. From the teacher perspective, the grading is done in a heartbeat. (Mrs. Black: post-observation interview)

Rather than describing her assessment of students in terms of what the students have learned and what she has learned about their level of understanding, this comment focuses on the grades for students. It also places the importance of student responses on whether they are correct or not, rather than what the responses tell the teacher and student
about what the student is thinking. Although students are receiving feedback on their performance, the feedback is in the form of a score or the correctness of individual responses and not a description of the quality of their work.

Awareness of student learning is one of the major components of rich formative assessment practice. Teachers who engage in rich formative assessment practice know about student learning and can identify specific learning difficulties. Conversely, teachers who focus on grades and desired textbook responses, or who have vague notions of student understanding, are not engaging in rich formative assessment practice. This continuum is summarized in Figure 5.4.

Figure 5.4. Awareness of student learning contributes to rich formative assessment

<table>
<thead>
<tr>
<th>Richness of Practice</th>
<th>Emerging Practice</th>
<th>Exemplary Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Awareness of Student Understanding (Assertion 4)</td>
<td>Focus on grades and summative assessment</td>
<td>Teacher has sense of whether the class “gets it” or not</td>
</tr>
</tbody>
</table>

**Assertion Five: Instructional Decisions are Based on Learning Evidence**

The last component in most descriptions of formative assessment is that the teacher uses data collected about student understanding to guide the choices he or she makes about subsequent instructional activities. These decisions could relate to more immediate decisions, such as the direction of the current class period, or other short-term decisions, such as those that affect the next few days of lessons. The formative assessment process could also influence longer-term decisions affecting the remainder of
a given unit of study, those related to the implementation of the course curriculum in the remainder of the school year, or even for future students in subsequent academic terms. Independent of the timing of the instructional decision-making process, teachers may make four types of decisions, each leading to a different degree of formative assessment practice. The analysis of teacher decision-making practices presented in this assertion addresses the focus question: *How do teachers use formative assessment data to make instructional decisions?*

*Teachers make individualized instructional decisions to meet each student’s learning needs.* The richest formative assessment practice comes from a teacher using data collected about student learning to individualize instruction and remediation based on a student’s observed learning needs. This requires the teacher to have knowledge of individual student performance. In a number of cases, teachers use the technology to identify specific learning needs of individual students. This facilitates their providing appropriate reinforcement and additional instruction. For example, students in Mrs. Black’s class describe how their teacher uses connected classroom technology to find out what each one of them knows in response to question prompts:

> She asks a question … if you do not know the right answer, she can click on [the incorrect answer] and figure out who it is and tell whether you are having problems. (student focus group; Mrs. Black’s classes)

Similarly, Mrs. Roberts tells the interviewer “I can see who is done first, and to get them onto something else, or I could go and help somebody that needed help” (Mrs. Roberts: post-observation interview). Such statements indicate that participants valued the ability to maximize their limited instructional time with students, accelerating the pace when warranted and providing additional reinforcement when necessary. Connected classroom
technology facilitates individualizing instruction through rapid aggregation of student responses so that the teacher can readily identify groups of students needing specific assistance.

It should be noted that in the current study, no assessment episodes related to individualized remediation are presented. The assessment episodes described here come from whole-group instruction, which is not conducive to individualized attention. Teachers may have used particular student responses to provide individualized instruction during the small-group portions of the class sessions observed. However, as described in Chapter 4, the design of the CCMS study was not such that conversations between the teacher and a small group of students were captured with high fidelity.

The teacher reinforces specific concepts with the whole class. Individualized and differentiated instruction poses significant challenges for implementation in secondary science classrooms. Because of these challenges, teachers may choose to not individualize instruction but rather to remediate particular learning goals with the entire class. This might take the form of posing additional problems for the class to solve, an extended discussion on a specific topic, or additional activities intended to guide students to a richer understanding of the content. The teacher is using his or her knowledge of the overall class performance to decide whether to move on to a new topic or to continue to review. Students in Mrs. Woods’s class tell an interviewer of a particular episode when the teacher used the data she collected about student understanding to adapt instruction accordingly.
Tuesday we were going over gravitational forces on the earth. Most of the people got right that it was 9.8 [m/s²] acceleration on the surface of the earth, and then the next questions were, what about the center of the earth. A lot of people still said 9.8, and it turned out to be 0. Then she realized that so many people got it wrong, and so she went over it more in depth than she had before. Now, I think everyone got it right on the quiz. (student focus group; Mrs. Woods’ classes)

Here, the teacher used responses to a question to judge that not enough students in the class understood that the acceleration due to gravity depended on the distance from the center of the earth. This led to her decision to review the concept in greater depth.

In another example, Mrs. Roberts uses connected classroom technology to uncover a confusing aspect of position-time graphs. She explains her realization and decision to adapt the next day’s lesson in a post-observation interview:

One thing that I noticed was that, from what I saw with the Quick Polls is that if the reference point or the starting point isn't right when the time starts, [the students] don't understand how the graph is going to change. So that's something we will have to review tomorrow. We can use the CBR, and whether they do it in groups, or we do it as a demonstration, then they need to see that just because the time starts, the object doesn't have to start moving. And I saw, from the Quick Poll, that many of them did not get that. (Mrs. Roberts: post-observation interview)

Her use of connected classroom technology to receive and aggregate responses from the entire class revealed that many students had difficulty interpreting a position-time graph for an object that did not start its motion immediately. Mrs. Roberts’s comments indicate that she is beginning to formulate a revised lesson plan for the next day’s lesson, incorporating specific activities that aim to clarify this particular point of confusion. This particular episode is one of the best examples of formative assessment identified from the data collected for this dissertation.

Sometimes, the information that teachers collect about student learning do not appear to lead to changes in instruction. One reason that this might be is that the teacher
has already planned lessons that take into account common student misconceptions. One of Mrs. Woods’s students comments on his perspective on this scenario:

“Lots of times, before she even turns up an answer, before she knows what people respond with, she already goes over the part that everyone ended up missing. She’ll say that she is sure a lot of people missed this, and then she teaches us, and then she shows it, and a lot of people did miss it.” (student focus group; Mrs. Woods’ classes)

This teacher has over 18 years of experience teaching high-school physics and physical science. With years of accumulated data on student learning, it appears that this teacher has already taken into account the most prominent student misconceptions and prepares her lectures and class discussions accordingly.

A similar situation arises when a teacher is responsible for multiple sections of the same course. Mrs. Woods relates a specific example of how what she learns in the first class period affects what she does in the third class period.

“Third period, I don’t think I did the exact same questions. I remember thinking that we probably don’t have to do this one again. I even did a second one – I did a question another time that as soon as I asked them, I thought that they didn’t need to answer the question, they already know this. If they know what a crest is, they know what a trough is… I guess what I do, is that I have this big list of questions that I just pick and choose from as I went down, depending on what I thought they needed to know.” (Mrs. Woods: post-observation interview)

Based on student performance on questions related to part of a transverse wave, she realizes that she could have streamlined her instruction more.

This experience is shared by other teachers as well. Mrs. Roberts discusses whether she adapts instruction during lessons with the interviewer.

Interviewer: during any of those [class periods] that I observed, was there a point where you had changed the course of the lesson because you had gotten something from students?

Mrs. Roberts: I didn't change it in the class that you observed, no, because I don't think that that particular example came up. Maybe, when I teach the same
thing period one, and then period three, I feel every day I do better in period three as a teacher because I can already know what is going to happen…. And often times, the next day I will start with them saying, "yesterday this worked better," and we go back and do something. (Mrs. Roberts: post-observation interview)

Clearly, Mrs. Roberts uses information gathered from one group of students to adjust her instruction for successive groups of students.

*The teacher describes vague changes in instructional strategies.* Although instructional decisions are made regarding specific content areas under study, not all teachers speak of their instructional decisions in specific terms. The vagueness with which teachers describe their intent to “do more review” may reflect an assumption that this refers to the content under study at the time. However, it may also reflect a lack of understanding of specific misconceptions held by students. Without knowing what specific learning difficulties are experienced by students, the teacher cannot provide appropriate remediation. Therefore, comments of this type are considered here to be a hallmark of relatively emergent formative assessment practice.

One example of a teacher expressing a vague notion of adapting instruction occurs in this description of a participant who is uncomfortable with one of the applications of the connected classroom technology. Class Analysis is a software program that allows the teacher to collect and score electronic files containing responses to a number of questions that students have answered on their handheld devices.

Interviewer: Which [software] components are you least comfortable using?  
Mr. Daniels: Well I guess the Class Analysis right now is. I really want to start learning how to use that more effectively and monitoring their progress with the content. (Mr. Daniels: spring telephone interview)
One might presume from the preceding comment that Mr. Daniels intends to follow student progress in order to identify struggling students or class-wide misconceptions in order to remedy them. However, this is not explicitly mentioned, and the conversation does not suggest that the teacher is equipped to follow through on this desire.

This may be contrasted with Mrs. Roberts’ description of how she does more reviewing and reteaching based on student input.

During the class period, of course if I Quick Poll them and I see that everybody’s completely off base and not where I want them to be, then I am going to backtrack and deviate from whatever my previous plan was. So that I can make sure that I retouch on something; to make sure that everybody has got it and I feel like we are good and then we go on. (Mrs. Roberts: spring telephone interview)

Although this conversation is general, content-wise, it does indicate that the teacher has a sense of how she would use the information that she collects about student learning. This indicates a higher level of formative assessment practice.

The teacher does not describe a change in instruction. In some cases, teachers may engage in no change to their instruction based on collected data about student learning. This may occur for a number of reasons. For instance, teachers may not have collected quality data on which to base their decisions. They may also feel constrained by time and curricular demands to cover content in a particular timeframe. No change in instruction may also reflect the tacit assumption by the teacher that students have attained a desired level of competency. For instance, Mrs. Roberts changes her approach to questioning students about correct and incorrect responses during a class wide discussion of responses to a mathematical problem related to velocity. Early in the discussion, Mrs. Roberts engages in the following five-iteration QRF cycle with her students.
Teacher: (Q₁) What about this one?
Students: (R₁) No. Wrong.
Teacher: (F₁; implied) (Q₂) Why?
Student: (R₂) They didn’t ask for distance.
Teacher: (F₂) They didn’t ask for distance; (Q₃) what did they ask for?
Student: (R₃) You’re supposed to average it.
Teacher: (F₃) So (Q₄) how should an average speed label look?
Student: (R₄) Miles per hour.
Teacher: (F₄) Miles per hour. (Q₅) How did you know to use miles and not kilometers?
Students: (R₅) It said miles. It said 500 miles.
Teacher: (F₅) It said 500 miles, good (Mrs. Roberts; Per. 1 Day 1)

Contrast this with a single-iteration QRF cycles occurring just a few moments later, where Mrs. Roberts and her students are discussing a different student response to the same question prompt.

Teacher: (Q) All right, this one; what’s wrong with this?
Student: (R) It’s wrong.
Teacher: (F) The math isn’t correct, ok? All right, I’m going to give you one more, and then we’re going to go on. (Mrs. Roberts; Per. 1 Day 1)

Mrs. Roberts is evidently convinced that by this point in the conversation that enough students have a sufficient grasp of the material that she is warranted in moving quickly through the remaining student responses. Note, however, that she is planning to administer an additional problem to students for them to solve, as an additional check for understanding.

A challenge in conducting studies on formative assessment arises at this point is the difficulty of knowing how the teacher would have taught the lesson without the data regarding student understanding. In the episode with Mrs. Roberts described just above, the second speed problem for students to solve was planned. Although she displays strong formative assessment practice in discussing right and wrong answers to a single prompt, carrying out the planned lesson regardless of how well students understand the
content is not a use of formative assessment data. This episode does not provide evidence of spontaneous formative assessment.

Interviews with other participants reveal a lack of formative assessment practice in their descriptions of the observed lessons. Mr. Daniels saw little to be surprised about in a class discussion where students were following along with class data and volunteering responses to formula-based calculations.

Interviewer: Yesterday in the class when the students were doing their calculations, were you surprised at the responses that they gave? They were calculating power and heat energy.
Mr. Daniels: Yes, heat energy and electrical energy using power and current.
Interviewer: Is it about what you expected?
Mr. Daniels: Yes. (Mr. Daniels: post-observation interview)

An inspection of the portion of the lesson being referenced in this discussion reveals that the teacher is asking students to solve problems and submit responses electronically. The teacher identifies a correct response and calls on a student to explain his reasoning in the following QRF cycle.

Teacher: (Q) Is 12,600 correct? Guys? Who put in 12,600? How did you do it Chris?
Male Student: (R) [inaudible]
Teacher: (F) Times 15. 12,600. (Mr. Daniels; Per 1 Day 1)

In this episode, several different student responses were submitted, but the teacher chooses to focus on the correct response. This allows him to have a student explain the reasoning behind arriving at the correct response. However, the focus on the correct answers means that Mr. Daniels misses the opportunity to explore student misconceptions and thereby an opportunity to carry out rich formative assessment practice.
A post-observation interview with Mrs. Black also fails to identify strong formative assessment. In this excerpt, the teacher is asked directly about any changes in instruction that were made during the observed lessons.

Interviewer: Is there any point of time during the lessons that you changed your plans to do things differently?

Mrs. Black: Nope. The only change I've made so far today is just that with the Screen Capture I was using, just clicking that login only button instead of having to scroll up and down through the list of everybody. Oh, and I did change, instead of doing the prediction for every substance, I only had them do it three times yesterday and then twice today, because I think they start to get bored with it every time. So we did it [together] a couple of times and then I let them do it by themselves. (Mrs. Black: post-observation interview)

Mrs. Black does indicate two changes that she made during instruction, although she initially does not consider to have adapted instruction at all. The first change she made relates to a technological adaptation that limits the number of blank calculator screens that appear when she uses Screen Capture, a snapshot-type functionality. The second change described is to streamline the process of having students work as a whole class to predict pH values of common substances then test them to verify those predictions. Neither of these changes relate to student understanding directly, but are connected to classroom management. The interview continues with a related question:

Interviewer: How might what you learned today about students’ knowledge change what you will do in subsequent lessons?

Mrs. Black: This is the first time I have used Screen Capture with [the students]. A lot of them were confused on exiting the Navigator and yet Navigator still working. It did take, with a couple of the periods, a bit longer for them to become adjusted to that. So I think since I've already run through it, I can go a little bit smoother in the future, it won't take much time. (Mrs. Black: post-observation interview)
When asked how she might adapt a future lesson based on what she now knows about student understanding, Mrs. Black responds with a discussion of classroom management and technological support that may enhance instruction, rather than a discussion of conceptual difficulties that students may be experiencing.

The types of instructional decision-making that teachers carry out based on data they collect about their students’ learning can therefore be summarized as in Figure 5.5. While teachers with strong formative assessment skills are able to use data to individualize and differentiate instruction or to provide additional whole-class instruction to address specific learning concerns, teachers whose skills are developing may be more likely to stick to an initial instructional plan or engage in general teacher-driven review.

<table>
<thead>
<tr>
<th>Richness of Practice</th>
<th>Emerging Practice</th>
<th>Exemplary Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Use of Data to Inform Teaching Practices (Assertion 5)</td>
<td>No changes in instruction based on assessment data</td>
<td>Vague “review and reteach” comments</td>
</tr>
</tbody>
</table>

Figure 5.5. Instructional decision-making

Summary

This chapter has provided interview evidence and examples of classroom practice that support the creation of a new fine-grained model of formative assessment practice, as outlined in the guiding question of this dissertation. Individual assertions addressed the focus questions; specifically, the assertions address the following components of formative assessment outlined at the beginning of this chapter:
1. What instructional tasks provide rich assessment information?

2. How can classroom-based assessment episodes be analyzed in detail while retaining the context of the related science instruction?

3. What characteristics of assessment episodes support rich formative assessment practice?
   a. What types of questions are used at the beginning of assessment episodes?
   b. What types of student responses are elicited?
   c. What strategies are used in subsequent questions to probe student understanding?

4. How do teachers gather data about student learning?

5. How do teachers use formative assessment data to make instructional decisions?

Furthermore, this chapter presented evidence for how connected classroom technology supports these aspects of quality formative assessment. Students can be more interactive with the content. Evidence of student learning is more apparent to both teacher and students. When each student responds, the teacher has a mechanism for identifying individual student misunderstandings, which can lead to improved instruction.

The model of formative assessment and connected classroom technology shown in Figure 5.6 presents significant implications for a variety of stakeholders; these will be discussed in the Chapter 6. Limitations of the current study will be presented along with recommendations for enhancements to a similar study. Finally, the next chapter will close with a discussion of future work using this model of formative assessment and how it can be used to improve the teaching and learning of secondary science.
### Richness of Practice

<table>
<thead>
<tr>
<th>How Instructional Tasks Are Used (Assertion 1)</th>
<th>Emerging Practice</th>
<th>Exemplary Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavior management, Grading</td>
<td>Teacher-centered content delivery</td>
<td>Student-centered, discussion-based</td>
</tr>
</tbody>
</table>

| Degrees of QRF Nesting (Assertion 2) | Single-degree QRF only | Some engagement in deeper probing; several degrees of QRF employed | A line of questioning is followed thoroughly, many degrees of QRF |

| Qualities of Initial Questions (Assertion 3) | Rhetorical or Knowledge | Comprehension | Application | High-order (includes Analysis, Evaluation, Synthesis) |

| Qualities of Student Responses (Assertion 3) | None, off-topic, or tangential | Don’t know, request to repeat question, yes/no response | Word or phrase | Hypothesis, conjecture, explanation |

| Qualities of Subsequent Questions (Assertion 3) | Implied or Leading | Repeat verbatim, ask student to repeat | Rephrase to clarify initial question | Request for elaboration |

| Awareness of Student Understanding (Assertion 4) | Focus on grades and summative assessment | Teacher has sense of whether the class “gets it” or not | Teacher can identify specific misconceptions held by the class | Teacher has evidence for each student’s level of understanding |

| Use of Data to Inform Teaching Practices (Assertion 5) | No changes in instruction based on assessment data | Vague “review and reteach” comments | Reinforce concept based on class-wide responses | Individualized remediation tailored to assessed learning needs |

* QRF refers to a cycle of Questioning, Responding, and Following-Up

---

Figure 5.6. Model of rich formative assessment practice in whole-class discourse
Chapter 6 : Discussion

This dissertation has provided a detailed analysis of formative assessment as practiced by teachers in middle and high school science classrooms equipped with connected classroom technology. This analysis resulted in the creation of a new model for rich formative assessment. The model presented in this dissertation is grounded in existing literature regarding formative assessment models but also emergent from classroom observations and interviews with classroom teachers and their students.

In this chapter, the overall model of formative assessment is summarized and aligned with the literature. Ways in which it extends current thinking about formative assessment practices and formative assessment models are also addressed. The influence of connected classroom technology on implementation of formative assessment strategies is then described.

This study is significant in that it provides both a methodology for conducting a detailed analysis of assessment episodes as well as in the creation of a formative assessment model. This significance is discussed in this chapter, along with a number of limitations to this study. The implications for practice with respect to a variety of stakeholders is then presented. This chapter concludes with a discussion of future directions for this particular body of work as well as a description of future studies that may arise from it.
Summary of Results

The previous chapter presented the result of this dissertation, culminating in a new model for formative assessment practice. In this section, the results are summarized and aligned with the current literature regarding formative assessment. Ways in which the results of this dissertation study extend the current literature on formative assessment are described. This section then discusses the role of connected classroom technology in supporting formative assessment as described in the model presented in this dissertation.

Comparison of Model with Existing Models

Figure 6.1 summarizes the formative assessment model described in Chapter 5. Each layer of this model presents a component of formative assessment with several benchmarks for levels of performance. In this section, the layers of this model are compared to current literature regarding formative assessment and aligned with existing models. The work in this dissertation synthesizes, complements, and extends the models presented by a number of other researchers (see, for example, Bell & Cowie, 2001; Black et al., 2004; Ruiz-Primo & Furtak, 2007; Torrance & Pryor, 1998).

Assertion 1: Implementation of high-quality instructional tasks. In the model presented by Torrance and Pryor (1998), the communication of an assessment task and the criteria by which it would be evaluated is the first phase of formative assessment. Similarly, Bell and Cowie (2001) describe the importance of eliciting information from students in planned formative assessment, and of noticing what students do while working on assessment tasks as part of unplanned formative assessment.
<table>
<thead>
<tr>
<th>Richness of Practice</th>
<th>Emerging Practice</th>
<th>Exemplary Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>How Instructional Tasks Are Used (Assertion 1)</strong></td>
<td>Behavior management, Grading</td>
<td>Teacher-centered; content delivery</td>
</tr>
<tr>
<td><strong>Degrees of QRF Nesting (Assertion 2)</strong></td>
<td>Single-degree QRF only</td>
<td>Some engagement in deeper probing; several degrees of QRF employed</td>
</tr>
<tr>
<td><strong>Qualities of Initial Questions (Assertion 3)</strong></td>
<td>Rhetorical or Knowledge</td>
<td>Comprehension</td>
</tr>
<tr>
<td><strong>Qualities of Student Responses (Assertion 3)</strong></td>
<td>None, off-topic, or tangential</td>
<td>Don’t know, request to repeat question, yes/no response</td>
</tr>
<tr>
<td><strong>Qualities of Subsequent Questions (Assertion 3)</strong></td>
<td>Implied or Leading</td>
<td>Repeat verbatim, ask student to repeat</td>
</tr>
<tr>
<td><strong>Awareness of Student Understanding (Assertion 4)</strong></td>
<td>Focus on grades and summative assessment</td>
<td>Teacher has sense of whether the class “gets it” or not</td>
</tr>
<tr>
<td><strong>Use of Data to Inform Teaching Practices (Assertion 5)</strong></td>
<td>No changes in instruction based on assessment data</td>
<td>Vague “review and reteach” comments</td>
</tr>
</tbody>
</table>

*: QRF refers to a cycle of Questioning, Responding, and Following-Up

Figure 6.1. Model of formative assessment through whole-class discourse
The dissertation study presented here incorporates these notions of assessment tasks by describing the types of tasks that teachers administer to students in secondary science classrooms. It goes beyond these studies, however, in that it provides several benchmarks by which formative assessment tasks can be identified and classified. Different levels of tasks can be aligned to formative assessment goals.

Assertions 2 and 3: Questioning is deep and involves higher-order questioning cycles. Questioning is a central aspect of formative assessment. Black and others (2004) describe how wait time enhances the quality of student responses and elevates them from being recitation of recalled facts. The quality of the teacher’s questioning therefore helps to move assessment beyond monitoring or diagnostic assessment to true formative assessment. This parallels well with the model presented in this dissertation, where exemplary formative assessment practice is aligned with higher-order questions that require students to provide richer responses.

In the Torrance and Pryor (1998) model, questioning makes up the second phase of a formative assessment cycle. They define questions as being either divergent, encouraging individual student thinking and a variety of responses, or convergent, leading to one answer or line of reasoning (Torrance & Pryor, 1998). These classifications correspond to the higher and lower levels, respectively, of questions in the model presented in this dissertation.

Similarly, the Ruiz-Primo and Furtak model of formative assessment addresses the notion of conceptual and epistemic questions. This matches best with the component of initial questions, which in the model arising from this dissertation are aligned with Bloom’s taxonomy. Conceptual questions are those that rely on basic knowledge and
understanding (Ruiz-Primo & Furtak, 2006), lower orders of Bloom’s taxonomy.

Epistemic questions are those that address procedures, interpretation of data, and evidence of knowing (Ruiz-Primo & Furtak, 2006), which are higher-order questions as in the dissertation model. The model presented here, then, extends the understanding of conceptual and epistemic questions by sorting questions into additional categories. Furthermore, coupled with the notion of QRF degree, the classification of initial as well as subsequent questions provides a richer understanding of how to identify high-quality questions within assessment cycles.

The Ruiz-Primo and Furtak (2007) model also describes an assessment cycle consisting of the teacher asking a question, the student responding, and the teacher recognizing the response and taking appropriate action, in “ESRU” cycles. It does not address, however, how to interpret assessment cycles where the outcome is that the teacher asks a follow-up question, engaging in an assessment cycle dovetailed with the original cycle. In the dissertation study reported here, a method for clarifying nested assessment cycles is described. By identifying initial questions as well as subsequent questions, the researcher obtains a mechanism for analyzing how deeply a concept is probed before moving on to a different concept or task.

Assertion 4: Teachers know more about student learning. For rich formative assessment to occur, the teacher needs to be aware of what students do and do not know. In the Torrance and Pryor (1998) model, this awareness arises in large part from the questioning practices implemented by the teacher. Deep probing of student understanding by the teacher uncovers aspects of student thinking that might not be
readily apparent. Without strong questioning skills, the teacher may remain unaware of what a student understands.

Bell and Cowie (2001) also present the notion of teacher awareness of student learning. In planned formative assessment, teachers elicit and interpret student knowledge; similarly, in unplanned formative assessment, teachers must notice and recognize it. The processes of interpreting and recognizing, respectively, align with the model presented in this dissertation in that teachers are becoming explicitly aware of what students do and do not understand. To a certain extent, this aligns with Wiliam’s (2008) diagnostic assessment category, where teachers engage in assessment to identify what difficulties students are having with their coursework.

The degree to which teachers know about individual student learning needs is summarized in the model of formative assessment presented in this dissertation. By presenting specific descriptions and examples, this model clarifies types of teacher knowledge of student understanding. Including this component of teacher awareness in the model presented here reinforces its importance and makes explicit the link between questioning activities and changes in instruction.

Assertion 5: Teachers use evidence to guide instructional decision-making. The final phase in the formative assessment cycle is for teachers to use the information they have collected about student understanding to shape subsequent instruction. Bell and Cowie’s (2001) description of formative assessment emphasizes the need for the teacher to respond in some way to the displayed student learning. Similarly, Ruiz-Primo and Furtak (2006) describe the role of the teacher in using assessment data to help students master content. Wiliam (2006) reinforces the importance of adapting instruction
according to student understanding; assessment is not formative if the teacher proceeds with planned classroom activities regardless of expressed student needs. These models, however, do not give much information regarding the ways in which teachers might adapt instruction to increase student achievement and understanding.

The final component of rich formative assessment practice presented in the model shown in Figure 6.1 supports the criteria for formative assessment as described above by demonstrating that teachers use the data about student learning to guide the choices they make regarding follow-up instruction. Furthermore, this model provides benchmarks for describing the ways in which teachers adapt instruction based on formative assessment data. These benchmarks extend the concept of formative assessment to provide a better understanding of the methods to implement changes in instruction. The change in instruction requires the teacher to have the requisite pedagogical content knowledge to make these decisions appropriately. Although the use of connected classroom technology can provide more accurate data about student learning more quickly than the teacher can aggregate herself, it cannot draw conclusions about the next steps to take.

Overall, the model of formative assessment practice presented in this dissertation provides a synthesis of four formative assessment models reported in the literature review (Chapter 2). Specific points of comparison of the model described here are summarized in Table 6.1. Components of the Black and Wiliam (Black et al., 2004) model highlighted in the model presented in this dissertation include an emphasis on teacher questioning as well as the role of tailoring instruction to meet learners’ needs. This dissertation complements the work of Bell and Cowie (2001) regarding the three phases of planned and interactive formative assessment. Torrance and Pryor’s (1998)
description of formative assessment defines a cycle paralleling the one described in this
dissertation, comprising task selection, gathering of student learning evidence, and
awareness of student work. Finally, the ESRU cycles described by Ruiz-Primo and
Furtak (2006; Ruiz-Primo & Furtak, 2007) are related to the QRF cycles described in this
dissertation as well as the component of using data to inform instructional decision-
making. The model of formative assessment presented in this dissertation therefore
incorporates aspects of previously described models. However, this model also extends
previous models of formative assessment by presenting a comprehensive, fine-grained
analytic tool for formative assessment practice.

Table 6.1. Comparison of Formative Assessment Models

<table>
<thead>
<tr>
<th>Model Component</th>
<th>Black &amp; Wiliam</th>
<th>Bell &amp; Cowie</th>
<th>Torrance &amp; Pryor</th>
<th>Ruiz-Primo &amp; Furtak</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rich Instructional Task</td>
<td>Setting task definition and criteria</td>
<td>Convergent versus divergent questioning</td>
<td>Elicit information through questions,</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interacting with student about work</td>
<td>Student response, Response by teacher</td>
<td></td>
</tr>
<tr>
<td>Deep &amp; Probing Questioning</td>
<td>Questioning skill is essential; especially wait time</td>
<td>Planned: <em>elicit</em> stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interactive: <em>notice</em> stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Awareness of Student Learning</td>
<td>Planned: <em>interpretation</em> stage</td>
<td>Observe students and give feedback</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interactive: <em>recognize</em> stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change in Instruction</td>
<td>Cites as critical; not in a model</td>
<td>Planned: <em>action</em> stage</td>
<td><em>Use of information</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interactive: <em>respond</em> stage</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other Features</td>
<td>Classify assessment as monitoring, diagnostic, formative</td>
<td>Emphasis on enhanced learning</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
**Contributions of CCT to Rich Formative Assessment**

The teachers whose classrooms are included in this dissertation study were in the first year of implementation of a connected classroom technology that has been shown in other studies (Pape et al., 2008) to significantly increase student achievement in mathematics classrooms. Other studies (Irving et al., 2009) of classrooms implementing this particular technology suggest that formative assessment is a contributing factor in promoting learning gains. The connection between rich formative assessment practice and connected classroom technology therefore needs to be clarified. According to the model of formative assessment presented in Figure 6.1, the first component of rich formative assessment practice is the implementation of instructional tasks that allow the teacher to uncover student thinking. This is an area in which the use of connected classroom technology influences formative assessment greatly.

A common use of connected classroom technology is in administering summative assessments: quizzes and tests can be implemented through connected classroom technology, reducing the burden of grading on teachers. This is not the only way in which connected classroom technology impacts instruction in secondary science classrooms, however. The use of CCT does not, in itself, bring about formative assessment. It does, however, provide a medium through which teachers can deliver rich instructional tasks that allow them to find out what students know. A major advantage of connected classroom technology is that activities can be designed so that all students are required to respond, by submitting data points, revealing their calculation process, or by moving a pointer on an on-screen image, for example. This public display of student thinking allows the teacher, at a glance, to assess her students.
The next several components of the formative assessment model described in this dissertation center on the role of deep, probing questioning and elicitation of student responses. Connected classroom technology does not have the ability to change the quality of the questioning and responding, but it does provide the teacher with a means of obtaining responses from all students. Moreover, it allows the teacher to aggregate all student responses rapidly and accurately, which in turn facilitates her making meaning of those results.

Because responses from all students can be made visible to the teacher, either in an aggregated or disaggregated form, the teacher is more aware of student learning, as described in the sixth component of the formative assessment model presented in this dissertation. Rather than waiting to grade summative assessments or relying on inaccurate measures such as students’ facial expressions and body language to determine the level of student comprehension, the teacher can identify specific ways in which students err. Partnered with skillful questioning and task development from the teacher, connected classroom technology supports formative assessment by enhancing the reliability of the assessment data gathered and processed by the teacher.

The last component of the formative assessment model shown in Figure 6.1 is the change in teaching practice that teachers may implement based on assessment data. The use of questioning applications can quickly provide teachers with information about which students answer correctly; the teacher can then use this information to differentiate instruction appropriately.

In summary, connected classroom technology does not cause formative assessment. The individual teacher must contribute a great deal of pedagogical skill to
implement rich formative assessment. Connected classroom technology does support, however, the teacher in this endeavor by allowing for instructional tasks that encourage the display of thinking by all students. This allows the teacher to engage in questioning that uncovers student learning and informs the teacher’s instructional decision-making process.

Significance and Implications for Practice

Significance of This Study

This study was conducted in secondary science classrooms equipped with connected classroom technology. The findings, however, are not limited to technology-enriched classrooms. Previous work has suggested a link between the use of connected classroom technology and formative assessment (Irving et al., 2009; Pape et al., 2008). Because of this link, rich formative assessment practices is likely to be found in these classrooms. The findings from this dissertation may be applied to understanding how formative assessment is carried out in a wide variety of classrooms, both technology-rich and those with low levels of technology.

This study is significant in that it presents new methodologies for analyzing and interpreting assessment episodes. As described above in comparison with existing models of formative assessment, this model includes a way of describing the degree of questioning, or the extent to which questioning cycles are nested. This dissertation advances the knowledge base of formative assessment by providing a fine-grained analysis of formative assessment practice as it occurs in secondary science classrooms.
A second major significance to this study is the model that is developed from the analysis of classroom observations and interviews with teachers and students. This is a more detailed model than others that have been previously described, which allows for a more precise evaluation of formative assessment practice by a variety of stakeholders. This model can be used to assess strengths and weaknesses of teacher practice from a variety of perspectives, including the classroom teacher, building and district supervisors, professional development providers, and university faculty and supervisors.

*Implications for Practice*

This dissertation study has resulted in the development of a new model for formative assessment that takes as a starting point the work of several other models of formative assessment. The model presented here, however, goes beyond these in a number of ways. It takes the most practical aspects of these and extends the descriptions of formative assessment. The fine-grained nature of the analysis present in this new model provides sharper focus to the components of formative assessment. Practitioners will find this model more understandable and resonant with classroom practice. The facets of the formative assessment model have a number of implications for a variety of stakeholders.

*How might students benefit from this work?* The purpose of formative assessment is to increase student learning. Therefore, for a model of formative assessment to be pedagogically relevant, it must be clear how it would benefit students. Indirectly, students may benefit from the model of formative assessment presented here because of the way in which it focuses the teacher’s attention on identifying misconceptions and tailoring instruction accordingly. Formative assessment should be, by its nature, a
student-centered classroom activity and this model promotes student interaction and engagement.

More directly, students may benefit from this work, particularly in the discussion of student responses. By identifying examples of student responses that provide richer formative assessment information to teachers, this dissertation indicates ways in which student responses can be enhanced. As teachers become more aware of how their questions influence the thoroughness of responses and adjust their questioning styles accordingly, students will be given more opportunities to communicate their thinking. Teachers then have more accurate information to base their instructional decisions on, thereby improving instruction and ultimately student learning.

How might teachers benefit from this work? Many classroom teachers are interested in continually improving aspects of their practice, and increasing student learning in particular. The model presented as part of this dissertation can assist classroom teachers in improving their formative assessment skills, especially when used as a self-evaluation tool. A teacher might choose to consider where on the continuum of each aspect of formative assessment practice he or she is currently practicing the most. For example, a teacher might find that she is strong in asking probing questions, but does not follow through with adapting her instruction to meet identified needs of her learners.

The strength of this model comes, in part, from the dissection of formative assessment into smaller, more practical components, combined with concrete descriptions of levels of practice. This not only affirms where on the formative assessment continuum their practice lies, but also encouraging teachers to consider how even richer practice might be implemented.
How might administrators benefit from this work? A classroom teacher is faced with many competing demands in a given day. The reality of the school setting is that on certain days, the class must engage in summative assessments. On other days, students may be engaged in small-group laboratory work, and the amount of whole-group discourse will therefore be minimal. On other days, whole-group discussion about the topics being studied is a more expedient instructional method, so that each student has the same opportunity to learn the material. Because of the variety of instructional needs in any classroom, the use of this model for any kind of summative evaluation is strongly discouraged, particularly if applied to a single observation period.

Despite not being useful as a teacher assessment tool itself, administrators may find this to be a useful tool in engaging an entire staff in dialogue about effective formative assessment practices. The department head or school principal might encourage groups of teachers to engage in action research projects where they participate in self- or peer-evaluation, dialogue about their findings, and develop plans for improving practice. Administrators who wish to promote professional development centered on formative assessment will find this model to be useful in helping teachers to identify and appreciate rich formative assessment practice.

How might teacher preparation programs benefit from this work? The work presented in this dissertation provides teacher education programs with a useful tool to use in evaluating the development of pre-service teachers. Formative assessment is a key pedagogical practice; teacher preparation programs should be encouraging all future teachers to acquire skill in this area. The model of rich formative assessment practice presented in this dissertation provides teacher preparation programs with a rubric that can
be used to indicate areas where pre-service teachers are relatively strong or weak. This in turn can indicate areas of potential growth.

With the many challenges facing them, novice teachers may select a few instructional strategies that provide them with some assessment data, such as “exit slips,” “bell-ringers,” or weekly quizzes, believing that they are indeed practicing formative assessment. However, this dissertation reiterates that the implementation of individual strategies does not, in itself, comprise formative assessment as much as how the teacher uses instructional strategies to uncover student learning. The model presented here can be used to broaden the perspective of novice teachers, demonstrating the importance of rich questioning strategies and adjusting instruction based on student responses.

*How might educational researchers benefit from this work?* This dissertation presents a new methodology for analyzing classroom discourse. The first aspect of this methodology comes from the partitioning of classroom transcripts into individual assessment episodes. This advances understanding of assessment beyond the single initiation-response-evaluation (IRE) model of questioning by demonstrating how subsequent probing questions can be linked to the initiating question of an assessment episode. This allows the entire context of an assessment episode to be considered together.

The second part of the methodology that benefits educational researchers is in the assignment of codes and their alignment to levels of formative assessment practice. The categorization of question and response types is not unique to this study, but the work in this dissertation does place them in the context of varying levels of formative assessment. This provides the researcher with a tool for evaluating questioning practices as a basis of
comparison, such as in a longitudinal study of teacher development or in comparisons across a variety of settings.

Taken together, the methods used in this dissertation present novel techniques for analyzing and interpreting assessment-based discussion in secondary science classrooms.

*How might technology developers benefit from this work?* Chapter 2 of this dissertation described several different forms of classroom technology designed to promote formative assessment. Studies of such systems indicate that they hold promise for enhancing student learning in a variety of ways, and these technologies are not likely to disappear from classrooms. Rather, new forms of technology will continue to be developed. This dissertation serves to benefit technology developers by highlighting relatively strong and weak aspects of formative assessment. New products aimed at enhancing formative assessment practice should take into account how the technology supports or promotes richer formative assessment practice.

One drawback to the system implemented in the CCMS study arises when the teacher chooses to use a software application that presents summary graphs of the proportion of students providing a given response, where a single bar on a histogram is used for each possible response. In open-ended response situations, only student responses that are identical are aggregated and tallied together. This results in a large number of responses that differ only by, perhaps, the number of trailing zeroes in a calculation (incorrect in higher-level science courses but acceptable in introductory science courses) or by two transposed letters resulting in a misspelled but recognizable term. A system that could recognize similar responses and automatically aggregate them
would greatly increase the teacher’s ability to quickly identify student learning difficulties.

Another difficulty that has become apparent with this particular system relates to text entry. One drawback to a handheld graphing calculator is that it does not employ “pretty print,” as referred to by a number of teachers using the technology. The characters on a graphing calculator are all the same size and only upper-case letters are possible. Exponents do not appear as smaller-sized numerals raised above the midline. Teachers and students have adapted to these inherent difficulties by finding alternate methods of expressing their intended responses. Another problem with text entry is that the alphabetically-arranged calculator keys are clumsy to navigate when typing in a longer open-ended response. A keyboard that uses a QWERTY-style interface would be easier to master. Related to this, a text-entry interface that uses the word-recognition software in text-messaging applications would reduce the time to type in responses and increase the likelihood that students would provide richer responses. Word-recognition software uses linguistic algorithms to predict the most common words that might be intended, even after only two or three letters have been entered.

Limitations of the Current Study

In this section, several limitations to this dissertation study are described.

Limited Numbers of Observations

Teachers were observed for two consecutive teaching days each year. No data other than teacher self-report and student perceptions recorded in surveys and the focus group is available for other days of instruction. No conclusions, therefore, can be drawn
regarding the absence of any specific teacher behaviors; the teacher may engage in other pedagogical behaviors during other class periods. These observations provide but a small window into the instructional practices of the selected participants.

This study uses data from the first year of technology implementation in participants’ classrooms. In some cases, the technology did not arrive in teachers’ classrooms until late autumn, at the time when the fall telephone interviews were being conducted. This means that when the first interviews were conducted, several participants had just received the technology and were still getting it set up in their classrooms. Since they had not had an opportunity to implement the technology in their teaching yet, interviews did not focus on pedagogy. Early interviews tended to emphasize difficulties with setting up and using connected classroom technology. While other studies (Irving et al., 2009; Shirley, Irving, Sanalan, Pape, & Owens, 2009) have examined the practicality of connected classroom technology implementation in secondary science and mathematics classrooms, this dissertation study focused on the pedagogy.

*No Student Collaborative Groups Recorded*

One of the major sources of data in this study was recordings of classroom observations. As such, the available data are subject to the ability of recording devices to pick up accurately the relevant conversations occurring in the classroom. With a single video camera and microphone, whole-class instruction could be captured with reasonable fidelity. Difficulties arose, however, when student responses were barely audible and when the teacher worked with individuals and small groups of students.
Fidelity of recording is the primary reason that small-group and individual instruction were not included in this study. The analysis of classroom assessment episodes is therefore limited to whole-class instruction. Other studies could be designed to specifically investigate small-group formative assessment, perhaps using multiple recording devices, etc.

With a well-defined coding system and trained observers, on-the-spot evaluation may be a possibility for analyzing assessment episodes. This would avoid the need for recording equipment, and the observer could more easily be positioned to capture teacher and student comments. This technique might be more faithful to the actual classroom conversation, but may prove difficult to carry out in real-time, as it would require an observer to only tally assessment conversations and not attend to other aspects of the classroom setting. Furthermore, without a recording, researchers would no longer be able to review tapes to check for accuracy of coding. In addition, if a comment was missed and not coded immediately, it would be lost from the data set. Furthermore, one of the major contributions of this work lies in the definition of assessment episodes that retain the context of science content instruction; to code conversations in real-time without capturing classroom conversations on tape would eliminate this context altogether.

Representation of Typical Teacher Practice

A major portion of this dissertation study focuses on the analysis of transcripts of classroom observations. The classroom observations were carried out with advance planning, which means that the teacher may have been demonstrating exceptional practice, rather than his or her normal use of the technology. Questions posed as part of
the student focus groups can serve to confirm teacher use of the technology, mitigating but not removing this concern.

Teachers are likely to want to present what they perceive to be their best teaching, and they know that the observer wishes to see connected classroom technology in use. This may cause the lessons taught during the observation days to be artificial and not reflective of that teacher’s typical classroom practice. The teacher surveys and interviews in this study serve as a self-report of practice. Participant self-report can be unreliable. Again, the use of several data sources is a way of triangulating the data to increase its trustworthiness, but care should be taken to not extrapolate any findings farther than the study warrants.

Limitations of the Data Set

The limitation discussed in this section is linked to the question of representative practice. From the participant’s perspective, many competing demands are being juggled when a researcher from the CCMS project arrives to conduct a two-day classroom observation. In an effort to impress the observer, teachers may allow classroom disruptions to go unchecked, so that the teacher is not recorded during a student confrontation. Furthermore, the attempt to display connected classroom technology use in as many forms as possible may also interfere with the participant’s ability to engage students thoroughly in a deep, rich learning activity.

In addition, the goal of the CCMS observations is to collect a broad range of classroom observation data, not specifically formative assessment data. Therefore, teachers are not asked to engage in formative assessment per se. Similarly, the classroom observation protocol is not refined for collecting formative assessment data, including the
design of interviews as well as recording of classroom interactions. Formative assessment is a major focus of the CCMS research project from which the data for this dissertation were taken, but it was not the only stated purpose of the study. Other purposes for the CCMS study include student engagement and motivation, the use of multiple representations of information, and self-regulated learning. The breadth of the CCMS project may therefore not be the best-suited for a deep and detailed analysis of formative assessment, or indeed of any one particular aspect of pedagogy.

This lack of specificity may be an advantage to the dissertation study presented here, however. If teachers were recruited for a study focusing only on formative assessment, they could have made very different decisions regarding how to implement connected classroom technology. In fact, teachers may have chosen to implement less creative uses of the technology, thinking that they were expected to demonstrate the quizzing and polling aspects of the TI-Navigator™ system. The design of the CCMS study is therefore advantageous in that no one particular focus is emphasized; the ability of teachers to prepare exactly what the researcher wishes to observe is limited.

**Demonstration of Student Achievement**

Formative assessment is formative when it leads to increased student understanding. Although this study examines elements of formative assessment cycles, no objective measures of student achievement are reported. Student pre-test and post-test achievement data are available for some students in the secondary science classroom included in this dissertation study. However, the small, pilot sample size, variation in the time of year the pretest was administered, and a wide range of grade and age levels
precludes using these results to make any claims regarding student achievement in these secondary science classrooms.

Student achievement data are available, however, for the mathematics portion of the CCMS research project. In the first year of implementation of connected classroom technology, gains in student achievement measured with an Algebra I pretest/posttest were significantly higher in classrooms using connected classroom technology than control classroom without CCT (Pape et al., 2008).

A wider question arises in considering whether to include measures of student achievement, and how to obtain these. Tests are not always a reliable measure of student achievement or understanding. Any assessment will bring with it specific limitations of its own, perhaps through a lack of cultural relevance or poor alignment with curricular objectives. The model presented here provides a means of evaluating formative assessment practice without incorporating external measures of student achievement.

*Identifying Changes in Instruction*

Another key component of formative assessment is that a change in instruction must occur, linked to the information that the teacher has learned about student understanding. This poses a significant challenge for the formative assessment researcher: how to know what a teacher would have done without the information?

Certain methods might be useful for examining changes in teaching practice. In a study focused on collecting data about formative assessment, teachers might be asked to prepare detailed lesson plans. Variances from the lesson plan could then form the basis of a semi-structured post-observation interview. To supplement this procedure, the participating teacher might be asked to view a videotape of the lesson along with the
researcher in order to describe the decision points in the lesson and recreate his or her thinking about how to adapt the lesson.

In classrooms where the teacher instructs multiple sections of the same course in a given day, comparisons may be made between sections. This would give evidence for how the teacher uses data gathered in one class period to adjust instruction in subsequent class periods. Interview evidence in the current dissertation study suggests that teachers may be using this kind of information to inform their instructional decisions in later class periods throughout the day; this may be a focus of future work in this area.

Without these methods having been utilized, no first-hand evidence for a real change in instructional practice can be ascertained. Assertion Five of this dissertation presents interview evidence with teachers, whose responses indicate that they do consider what they have learned from one class period in proceeding through the rest of the day’s lessons.

Limitations of Modeling

In this dissertation, a conceptual model of formative assessment practice is presented. Conceptual models can be described as “mental constructs consisting of symbols, images, and propositions that together characterize a certain category of events” (Gilbert & Ireton, 2003, p. 8). The model described in this dissertation uses a graphic means to characterize the events that comprise rich formative assessment in secondary science classrooms. It includes delineations of events (signified by aspects of practice described as exemplary) as well as non-events (signified by emerging practice).

It is important to remember that a model is merely a representation of actual events and behaviors and must be interpreted accordingly. The model here is subject to
the researcher’s own biases and experiences and is not a perfect depiction of formative assessment. Moreover, much work remains to be done to define more clearly the components of formative assessment practice outlined in this model and to elucidate the causal relationships between each element and a hypothesized increase in student achievement.

Despite these shortcomings, the model generated through this dissertation study marks a step forward in the understanding of the process of formative assessment. Gilbert and Ireton (2003) set criteria that models must be transparent in their representation of the target, related to other models, robust to changes in assumptions, fertile in providing new avenues of inquiry, and easy to enrich. The model of formative assessment presented here fulfils each of these criteria.

Future Work

Changes in Formative Assessment Practice

The work presented in this dissertation should be extended in a number of different ways. Some of these extensions are refinements of the model and the methodology. One area of the model that should be refined is the category of student responses, “word/phrase.” This category was used to code 65% of all student responses analyzed here (see Table 5.7 and subsequent discussion). An additional categorization may be identified for this part of the coding scheme in order to elucidate further the kinds of responses that students are providing. For instance, a single word response (“acid” or “velocity”) might be part of one new category, perhaps called *vocabulary term*, while a longer phrase (“it goes up”) might be termed *describe event*. Ideally, a refinement of this
category would not involve any more than two additional categories to keep the coding system from becoming too unwieldy, but an additional level of detail may provide more insight into the kinds of responses that students provide during class discussions.

This model is based on the teaching practices of four secondary teachers who participated in a broader research study regarding the implementation of technology. Although the teachers’ instructional contexts differ greatly, this model is still based on a limited sample. Future work should be directed at verifying this model’s applicability to other classroom contexts. Specific instructional settings to identify include classrooms of other disciplines, such as mathematics, social studies, and language arts, as well as those that do not employ connected classroom technologies.

Furthermore, the role of connected classroom technology in supporting formative assessment should be investigated further. Is the critical component the use of technology to achieve the objectives of assigning rich, student-centered instructional tasks and in rapidly aggregating evidence of student learning? Or do other strategies not using computer-based technologies provide the same benefits? The model presented in this study would serve as a useful tool to assess the degree of formative assessment occurring in a variety of groups of classrooms, each with a different intervention that provides the features of a connected classroom.

Another potential aspect of this model that is not addressed here will be termed “path.” The path of an assessment episode has to do with who initiates an assessment episode and a detailed accounting of who speaks at what times. For instance, an assessment episode might be coded as teacher-student-teacher (T-S-T), to illustrate a teacher asking an initial question, a single student responding, and the teacher providing
feedback to this response. Other possibilities include student-initiated assessment episodes (S-T-S) or discussions involving many students sharing ideas in turn before the teacher intervenes (T-S₁-S₂-S₃…Sₙ-T). In the current study, only two out of 294 assessment episodes would be termed S-T-S. All other episodes were of the T-S-T variety (data not shown). This is an interesting avenue of pursuit because of the ability of connected classroom technology to increase student engagement and community-centeredness (National Research Council, 2000).

Other extensions to this dissertation involve follow-up studies in which the model developed here will be applied to other studies of connected classrooms. The participants in this study were followed for three consecutive years of practice. Interview, survey, and classroom observation data, as described in Chapter 3, is available for these participants for all three years. This model of formative assessment practice may be a useful tool with which to interpret potential changes in assessment practice as the teachers become accustomed to implementing connected classroom technology into their science teaching.

Furthermore, the CCMS project from which the data analyzed here were obtained was designed as a randomized controlled trial study of Algebra I teachers. Therefore, there exist the same types of data (interviews, surveys, observations) as well as comparable pre- and post- achievement tests from a large number of mathematics teachers. This allows for the model presented in this dissertation to be applied to similar mathematics classrooms, first to test the model to determine its applicability to disciplines other than science, and then to apply it to analyzing teacher practice.
Other potentially rich avenues of exploration include comparisons of formative assessment practice throughout the day in classrooms where the teacher has multiple sections of the same course (same content and activities, different group of students). This would help to address how teachers adapt their instruction from class period to class period. Does questioning get richer, or does it get weaker as teachers perceive that they have already learned enough about their students’ thinking and assume that the difficulties experienced in the morning will also affect students in the afternoon?

Enhancing Formative Assessment Practice

This model also presents a number of rich opportunities for further research regarding teacher development. Professional development opportunities for teachers often focus on the sharing of teaching strategies; tips and tricks for engaging students in hands-on exploration. This focus on strategies overlooks the point being made with the model presented here, that formative assessment can be accomplished with a variety of instructional tasks but that the teacher must probe for student understanding and modify instruction accordingly. Teachers need assistance in developing these skills. The ways in which teachers implement the skills of questioning, recognizing student difficulties, and adapting classroom activities to ensure student learning are what results in increased achievement, not simply the administration of innovative instructional strategies and tasks.

A number of the benchmarks on the model presented in this dissertation highlight the need for additional research. First, research is needed in how teachers learn how to use questioning strategies. One way in which this might be accomplished is to have teachers record a class session and apply the analysis presented in Chapter 5 of this
dissertation to their work. Using the results of this analysis, teachers could then develop an action research project aimed at identifying an area of questioning practice to improve.

A critical area of new research based on this model is in examining how teachers use evidence of student learning to make instructional decisions. An interview-based protocol could be developed to find out what factors teachers employ in making their instructional decisions. Teachers could also be presented with scenarios to consider and explain how they would make instructional decisions. Teachers could also be asked to administer a formative assessment, perhaps in the form of an exit pass, and engage in reflection with an interviewer immediately after collecting the student responses.

Use of this model would demonstrate if a teacher were able to identify student learning needs but be unable to articulate ways to change instruction. Additional work is needed to clarify and communicate to teachers possible follow-up teaching strategies to help students master concepts. For example, when faced with students struggling with a science concept, teachers might choose to employ analogies, multiple representation of information, modeling, inquiry-based experimentation, or other strategies. Alternatively, teachers could benefit from guidance in how to reflect productively on their own practice to find ways to help students learn. The teacher may have unwittingly presented a series of sample problems that share a common feature, for instance, that lead students to develop algorithms that work for those examples but cannot be applied to problems that are more difficult. Work in this area would enhance formative assessment by guiding teachers to move beyond the “review and reteach” phase of changing instruction.

It should also be remembered that becoming an effective teacher is a developmental process. Novice teachers will not, and should not, be expected to perform
at the same levels as master teachers, nor even at the same levels as teachers with a few years of experience. An informative study would be to investigate the formative assessment practices and beliefs of teachers at a variety of stages of professional development in a quasi-longitudinal study. Moreover, a true longitudinal analysis of pre-service teachers and how they acquire formative assessment skills (through their teacher preparation program, entry year, first 2-5 years of teaching, and even after perhaps 10-15 years of teaching) would provide important insights into how teachers develop sophisticated pedagogical skills.


Appendix A: Teacher Demographic Survey
TI-Navigator Research Project
Demographic Information Form
Summer 2006-Cohort 3

1. Participant code: ____________________ 2. Today's Date ___/___/_______ Instrument Protocol 01


5. Number of years teaching: __________ 6. Number of years teaching Physical Science: __________

7. List courses you will be teaching during the 2006-2007 school year:

8. Anticipated Typical (Maximum) number of students in a typical class

9. Anticipated number of Physical Science sections

10. School start date ___/___/_______

11. Physical Science class end date ___/___/_______

11. Name of physical science textbook:

12. Ethnicity

☐ American Indian/Alaska Native

☐ Asian

☐ Black or African-American

☐ Hispanic or Latino

☐ Native Hawaiian or Pacific Islander

☐ White

☐ Other (please specify) ______________________

13. Gender

☐ Male ☐ Female

14. Undergraduate degree major: ____________________ Minor ____________________ year awarded __________

15. Graduate degree: ______________________________ year awarded __________

16. Other degrees:

17. Certification or licensure area:

18. Grade span for license:
18. What are your expectations for the TI navigator week-long institute?


20. What do you anticipate will be three impacts of using the TI-Navigator on your classroom?
   1)


   2)


   3)
Appendix B: Teacher Technology Use and Professional Development Survey
Technology Use and Professional Development Survey

Your Initials  Participant ID  Today’s Date: / /

Instructions: Please fill in and submit your responses.

<table>
<thead>
<tr>
<th>General Technology Use &amp; Skills</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>My ability to use computers to</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Compose / write papers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Assist in personal record keeping</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Deliver individual learning (computer aided learning)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Aid in statistical analysis and research</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Aid in class management (develop syllabi, track grades)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Entertain oneself (games)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Design instructional materials</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Teach students at a distance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9. Create images</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10. Manipulate images</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Educational Technology Use in Science Instruction</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>My ability to use educational technology in instruction for...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. Data collection (pH meters, temperature probes, motion probes etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12. Database storage of lab data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. Preparing graphs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14. Demonstrations and modeling</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15. Computer assisted instruction (CAI)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16. Problem solving</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17. Materials generation (problem sets, quizzes, tests, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18. Individualized instruction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19. Spread sheet analysis of collected data</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20. Internet resources (web sites)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21. Internet resources (E-mail)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## Educational Technology Use in Science Instruction

<table>
<thead>
<tr>
<th>Activity</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>My ability to use specific technologies for science instruction.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22. Graphing calculators</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23. Spreadsheets or Excel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24. Web page authoring software</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25. Digital cameras</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26. Power Point</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27. Simulations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28. Personal Digital Assistants (such as Palms)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>29. Digital microscopes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30. GIS (Geographical Information Systems; such as ArcExplorer, Google Earth, GeoData, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31. Virtual planetariums</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### What is your skill in using probes to measure...

<table>
<thead>
<tr>
<th>Skill</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>32. Temperature</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>33. Motion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>34. Force</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35. pH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36. Conductivity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>37. O2 level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38. Light</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39. Pressure</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## Skills

<table>
<thead>
<tr>
<th>Skill</th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>My ability to...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40. Save an image to a disk or to my computer.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41. Identify the file extension for an electronic file.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>42. Use a scanner.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43. Burn a CD.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44. Edit an electronic image.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>45. Include an image in a word document.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>46. Create tables in an electronic document.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47. Use concept mapping software.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48. Install Operating System on a graphing calculator</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49. Install software on a computer</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50. Set up a computer network</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>51. Write an original computer program.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Professional Development:

Below indicate your level of participation in each of the following within the last two years:

<table>
<thead>
<tr>
<th>Professional Development</th>
<th>A: Never</th>
<th>B: Once in two years</th>
<th>C: Once a year</th>
<th>D: Twice a year</th>
<th>E: More than twice a year</th>
</tr>
</thead>
<tbody>
<tr>
<td>52. Science content courses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>53. Physics Content Courses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>54. Chemistry Content Courses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>55. District level one-day workshops</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>56. Long-term professional development (2 or more days)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>57. Graphing calculator workshops</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>58. Technology workshops</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>59. University Course work in education/ pedagogy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>60. Attendance at a professional conference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>61. Presenting at a professional conference</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>62. Others: (Please specify)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Expertise

My level of expertise in:

<table>
<thead>
<tr>
<th>My level of expertise in...</th>
<th>A: Novice</th>
<th>B: Somewhat Adequate</th>
<th>C: Adequate</th>
<th>D: Somewhat knowledgeable</th>
<th>E: Very knowledgeable</th>
</tr>
</thead>
<tbody>
<tr>
<td>63. Newton’s Laws of Motion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>64. Momentum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>65. Work, power and energy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>66. Gravity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>67. Projectile and Satellite Motion</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>68. Heat and temperature</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69. Electricity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>70. Magnetism</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>71. Waves and sound</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>72. Light and optics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>73. Properties of matter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>74. The periodic table</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>75. Atoms and atomic theory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>76. Chemical equations</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>77. Mixtures</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>78. Chemical bonding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>79. Acids and bases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>80. Oxidation and reduction</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>81. Organic chemistry</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>82. Environmental science</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>83. Earth science</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>84. Biology</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>85. Graphing and graph interpretation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Appendix C: Fall Teacher Telephone Interview Protocol
Teacher Fall Telephone Interview Protocol, Year 1

Teacher ID: _______________  Date: _______________
Teacher Initials: _______________  Interviewer Name _______________

*Remember to ask about missing surveys or measures.*

1. I'm interested to hear about your successes and challenges in the connected classroom so far this fall.
   a. Successes?
   b. Challenges?

2. For what purposes did you use TI-Navigator this fall?
   a. Quick poll
   b. Activity center
   c. Learn check
   d. Class analysis
   e. Probe ware
   f. Study cards
   g. Homework check
   h. Other

3. How did your students respond to the connected classroom technology this year?
   a. Positive
   b. Negative
   c. Mixed
   d. other

   a. Not at all
   b. Some
   c. Greatly

5. What examples can you provide of how use of the connected classroom changed your teaching strategies or plans?

6. On a scale of 1 to 10, how would you rate your level of implementation of connected classroom technology?

7. What plans do you have for connected classroom use for the near future?

   Do you have any questions for me?
Appendix D: Spring Teacher Telephone Interview Protocol
Teacher Spring Telephone Interview Protocol, Year 1

Teacher ID: ____________________________ Date: ____________________________
Teacher Initials: ____________________________ Interviewer Name ____________________________

Remember to ask about missing surveys or measures.

1. Approximately when did you begin to use the TI Navigator system this school year?
2. What was the first day of instruction for the class that is part of our study? __________
   What was (will be) the last day of instruction for the class that is part of our study? ______
3. Approximately how often have you used the TI Navigator each week during the following time periods?

<table>
<thead>
<tr>
<th>Time Period</th>
<th>Estimate Frequency of TI Navigator use each week</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sept-Oct-Nov</td>
<td></td>
</tr>
<tr>
<td>Dec-Jan-Feb (before T-cubed conference)</td>
<td></td>
</tr>
<tr>
<td>March-April-May (since T-cubed conference)</td>
<td></td>
</tr>
</tbody>
</table>

4. I'm interested to hear about your successes and challenges in the connected classroom this year.
   a. Successes?
   b. Challenges?
   (Can you give an example? Can you elaborate further? Can you give another example?)

5. On a scale from 1 to 10, how would you describe your comfort level using the TI Navigator system at this time? Please provide a rationale and examples for your rating. [How has this changed from Year 1 (or from Year 2)?]

6. Which TI Navigator components do you feel most comfortable using? [How has this changed from Year 1 (or from Year 2)?] (Can you give an example? Can you elaborate further? Can you give another example?)

7. Which TI Navigator components do you feel least comfortable using? [How has this changed from Year 1 (or from Year 2)?] (Can you give an example? Can you elaborate further? Can you give another example?)

8. On a scale from 1 to 10, how would you rate your own implementation of the classroom connectivity technology? Please provide a rationale and examples to help me understand your rating. [How has this changed from Year 1 (or 2)?] (Can you give an example? Can you elaborate further? Can you give another example?)

continued
Appendix D: Spring Teacher Telephone Interview Protocol continued

9. Estimate how frequently and for what purposes you used each component of TI-Navigator since the beginning of school? [never, sometimes, often, every class period]
   a. Quick poll
   b. Learn check
   c. Activity Center
   d. Class analysis
   e. Screen capture
   f. Data collection/analysis
   g. Probe ware

10. Has the use of the TI Navigator supported your understanding of your students’ knowledge of Algebra I [physical science]? If so, then how has it supported your understanding? (Can you give an example? Can you elaborate further? Can you give another example?)

11. How have your students responded to the TI Navigator technology?
   a. Positive
   b. Negative
   c. Mixed
   d. Other
      (Can you give an example? Can you elaborate further? Can you give another example?)

12. Are there any major differences between the atmosphere in your connected classroom in this academic year, compared with classes in other years [before you started using TI-Navigator]? If so, what are they? (Can you give an example? Can you elaborate further? Can you give another example?)

13. How has using the TI Navigator technology influenced the way you plan and implement instruction? (Can you give an example of before instruction planning? Can you give an example of changes made during instruction? Can you elaborate further? Can you give another example?)

14. How has use of TI Navigator technology changed how you interact with your students in your classroom? (Can you give another example? Can you elaborate further? Can you give another example?)

15. On a scale from 1 to 10, how would you rate your implementation of the pedagogy of the connected classroom (SRL, formative assessment)? Please provide a rationale and examples to help me understand your rating. [How has this changed from Year 1 (or 2)?]
16. What plans do you have for using TI Navigator next year? What lessons will you repeat that you did this year? What new lessons do you plan to try? (Can you give an example? Can you elaborate further? Can you give another example?)

17. Have you made any presentations regarding use of the Navigator technology? Are you planning to present during the summer?

Do you have any questions for me?
Appendix E: Teacher Post-Observation Interview Protocol
Teacher Post-Observation Interview Protocol

Classroom Connectivity – Classroom Observation Protocol
Post-Observation Teacher Interview

Participant #: __________  Teacher’s Initials: __________  Date of observation: __________

NOTE: Instructions related to applicable cohorts are provided in square brackets. Terms in parentheses are intended for TI-Navigator users (i.e., cohorts 1-3 during years 2-4 and cohort 4 during years 3-4). Terms in brackets are for Graphing calculator users (i.e., cohort 4 during year 2).

1) Tell me about your plans for this lesson. Try to reconstruct your thought processes while planning this lesson. What decisions did you make while planning this lesson? Tell me about how you intended to use (the TI Navigator) (graphing calculators) when you planned for this lesson.

2) Tell me about how you have taught this lesson in the past.

3) How might you have done this differently if you were not using (the classroom connectivity technology) (graphing calculators)?

4) What changes did you make to plan this lesson (using TI Navigator) this year?

5) [For Cohort 1 during Year 2-4; for Cohorts 2 & 3 during Year 3 and 4; for Cohort 4 during year 4] Did you teach this lesson using the TI Navigator differently last year? How was it different this year from last year?

6) How did you use (the TI Navigator) (graphing calculators) specifically to support student learning?

7) During the course of teaching, did you change your plans for the lesson? How?

8) How might what you learned today about students’ knowledge change what you will do in the subsequent lessons?

NOTE: Additional questions will be formulated based on teacher’s responses and observation running record. The purpose of the teacher interview is to probe classroom observation deeply.

continued
Appendix E: Teacher Post-Observation Interview Protocol continued

Participant #: ________  Teacher’s Initials: ________  Date of observation: ________

9) <If observer has observed additional uses of the TI Navigator or graphing calculator, provide list of uses of (the TI Navigator) (graphing calculators) observed during class but not discussed previously in interview.>
   I noticed that you used (e.g., learn check) [the graphing calculator] when you were at X point in the lesson doing Y. Can you talk about how you used the technology at these times? How did you make the decision to use the technology at these times?

10) In general, how has using the classroom connectivity technology changed from Year 1 to Year 2?

11) Which TI Navigator components do you feel most comfortable using?
    Can you give an example? Can you elaborate further? Can you give another example?

12) Which TI Navigator components do you feel least comfortable using?
    Can you give an example? Can you elaborate further? Can you give another example?

13) On a scale of one to ten, how would you rate your implementation of the (Navigator technology) (graphing calculators) in your teaching this year?

14) [For Cohort 1-3 during Year 2-4; for Cohort 4 during Year 3 and 4]
    Considering our discussions of the pedagogy for using the TI Navigator, on a scale from one to ten how would you rate your implementation of the pedagogy at this point?
    [For Cohort 1 during Year 2-4; for Cohorts 2 & 3 during Year 3 and 4; for Cohort 4 during year 4]
    How has it changed from Year 1 (2 or 3) to Year 2 (3 or 4)?
    Can you give an example? Can you elaborate further? Can you give another example?

NOTE: Additional questions will be formulated based on teacher’s responses and observation running record. The purpose of the teacher interview is to probe classroom observation deeply.
Appendix F: Student Observation Survey
TI-Navigator Research Project Student Survey

For each of the following statements indicate whether you Strongly Disagree (SD), Disagree (D), Neither agree nor disagree (N), Agree (A), or Strongly Agree (SA) by circling the appropriate choice to the right of the statement.

1. Using the TI-Navigator does not help improve my understanding
   - SD
   - D
   - N
   - A
   - SA

2. Class dynamics are not affected by the use of the TI-Navigator
   - SD
   - D
   - N
   - A
   - SA

3. The teacher knows just as much about my understanding without the TI-Navigator as with it
   - SD
   - D
   - N
   - A
   - SA

4. There is a greater sense of community in a TI-Navigator class than in other classes
   - SD
   - D
   - N
   - A
   - SA

5. The TI-Navigator helps the teacher tell if I understand a concept
   - SD
   - D
   - N
   - A
   - SA

6. I am more actively engaged in a TI-Navigator class than in others
   - SD
   - D
   - N
   - A
   - SA

7. The TI-Navigator makes no difference to my effort in answering questions
   - SD
   - D
   - N
   - A
   - SA

8. I find no advantage in using the TI-Navigator to help me build on my knowledge
   - SD
   - D
   - N
   - A
   - SA

9. Some TI-Navigator questions make me try really hard to answer them
   - SD
   - D
   - N
   - A
   - SA

10. I am less on task in TI-Navigator classes than other classes
    - SD
    - D
    - N
    - A
    - SA

11. Using the TI-Navigator does not help in letting me know where I stand on a question
    - SD
    - D
    - N
    - A
    - SA

12. Using the TI-Navigator I can quickly tell whether or not I am right or wrong
    - SD
    - D
    - N
    - A
    - SA

13. Doing activities with the TI-Navigator in class helps me get a better understanding of concepts
    - SD
    - D
    - N
    - A
    - SA

14. Class interactions resulting from using the TI-Navigator help my learning
    - SD
    - D
    - N
    - A
    - SA

15. Doing activities in class with the TI-Navigator helps me relate new material to things I already know
    - SD
    - D
    - N
    - A
    - SA

16. Using the TI-Navigator does not improve the sense of community in classes
    - SD
    - D
    - N
    - A
    - SA
Appendix G: Student Focus Group Protocol
TI-Navigator Research Project

Student Focus Group Protocol

The student focus group will take place after the classroom observation. Its principal purpose is to elicit amplification on responses to the survey with an emphasis on self-regulated learning.

The focus group discussion should begin with the following questions relating to classroom use of the TI-Navigator:

- How typical was the use of the TI-Navigator in your lesson today?
- What do you think was the purpose of the use of the TI-Navigator today?
- For what other purposes has the TI-Navigator been used?
- What do you like or not like about the TI-Navigator system?

The focus group discussion should then be directed more deeply toward exploring students’ responses on survey items by asking the questions listed below. With the exception of #3, the second bulleted question in each group should only be used if necessary, i.e., if students don’t understand or don’t fully respond to the first bulleted request for elaboration.

(Survey statement 2)
1. What does it mean to be actively engaged in this class?
   - Tell me about the TI-Navigator and your classroom engagement.
   - Does the TI-Navigator change the way you engage in the class? How?

(Survey statement 3)
2. Do you try hard to answer questions in this class?
   - Tell me about the TI-Navigator and your effort in this class.
   - How does the TI-Navigator affect your effort to answer questions?

(Survey statement 1)
3. What helps you know whether you are doing well in this class?
   - Does the TI-Navigator help you know how well you are doing on learning a concept? How?
   - What kind of information does it provide that you find useful?
   - What do you do once you have information about how you are doing?

continued
Appendix G: Student Focus Group Protocol continued

(Survey statement 8)
4. What kinds of activities help you make connections between new information in class and things you already know in mathematics [science]?
   - Tell me about the TI-Navigator and the connections you make between new information in class and things you already know in mathematics [science].
   - Does the TI-Navigator help you make new connections to things you already know? How does it help?

(Survey statement 4)
5. How do you think the teacher knows when you understand a concept in this class?
   - Tell me about the TI-Navigator and your teacher’s knowledge of your understanding.
   - Does the TI Navigator help the teacher know when you understand a concept? How?

(Survey statement 5)
6. How do class interactions help your learning?
   - Tell me about the TI-Navigator and classroom interactions.
   - Are classroom interactions different in a TI-Navigator classroom than in other classrooms? How?
Appendix H: Coding Book
Appendix H: Coding Book

Identifying Instructional Tasks from interviews

1. Using interviews (FTIP, STIP, POI, SFG), read and identify descriptions of instructional tasks. These may or may not include Navigator. The description of an instructional task often includes the interviewer's question (for the sake of context). A single instructional task is a description of a teaching strategy (questioning students, collecting lab data, etc) or a plan for instructing a specific topic.

2. After identifying all mentions of instructional tasks, code by hypothetical task, general task, specific task, non-task.
   a. Non-examples: These include descriptions of playing calculator games and technology concerns.
   b. Hypothetical: These include descriptions of units of study or topics that the teacher is intending to teach later in the year or in future years. It might also include descriptions of how else the teacher might have taught the unit other than the observed strategies. Hypothetical instructional tasks also include descriptions of how the teacher intends to use other components of the Navigator system in the future or when s/he feels more comfortable.
   c. General: These are instructional tasks without the context of a specific topic of study or a specific component use.
      - "we used it for weather and oceans"
      - "I use it for weekly review questions"
      These are general tasks because the teacher has not described any of the instructional methods used. These are descriptions of overall methods or of units that may last for several days to several weeks.
   d. Specific: These can be specific in terms of the instructional strategy used (sending the students a "Do-now" followed by review of student responses and reteaching, for example) or the topic of study (displaying motion graphs created with CBRs). A specific task typically concerns an intact lesson, lasting perhaps 15 minutes to several days. After specific tasks are identified, they are classified according to the teacher and arranged by the topic area discussed.

continued
Appendix H: Coding Book continued

Defining Assessment Episodes from Observation Transcripts

1. Read through transcript to identify questioning cycles. Only consider events that are instructional; do not include classroom mgmt unless it disrupts the instructional sequence. Do not include materials management and similar instances where teacher is asking about progress or procedures not related to learning science content or science procedures. Skip descriptions of after-school activities and other chit-chat.

2. Limit assessment episodes to whole-class instruction

3. Limit to actual science-based content and process skills. Not classroom management and not procedures related to technology use, unless it is explicitly related to science learning.

4. Identify AEs based on one discrete question (from T or from S). It may or may not have other follow-up associated with it; this is part of what will be coded. Every question that begins a new set of probing starts a new AE. Demarcations include “now let’s look at this one,” moving to the next problem of the homework, etc.

5. An assessment episode lasts until there is a new probing question or some other event (non-instructional activity, activity without assessment-type discourse such as completing the activity) occurs to break it up.

6. Whether a Q belongs with a new AE or the existing AE has to do with how well it stands on its own. If it relates to that original Q, it is part of the AE. If it is all part of an assessment episode there is a (implied or explicit) overarching question/task being posed to students. If two cycles occur adjacent to one another, without teacher talk in between, it can be difficult to distinguish them as discrete cycles or as a nested cycle. If the teacher affirms a student response in a way that accepts it as the final answer to the previous question, this may indicate that they are two separate questions.

continued
Coding/Categorizing Parts of Assessment Episodes

1. Copy all references from the node for that observation transcript into Excel. Break up so that each part of the Q-R-F cycle is on a fresh line. Follow-up and Subsequent Question moves may be implicit; include blank lines when these are not identifiable.

2. Use additional columns for the following coding aspects:

3. Path: Use notation of “T,” “S,” “SS” to indicate whether the teacher (T), a single student (S), or multiple students (SS) have provided the utterance.

4. Part of Cycle: Following each line of the transcript, indicate whether it is a Question, Response, or Followup.

5. First cycle: designate as 1⁰ (QU1, R1, F1)

6. Follow-up events may be to initiate a new QRF cycle; indicate these as 2⁰, 3⁰, and so on. (QU1, R1, F1, QU2, R2, F2, and so on)

7. Quality Codes: Using the codes presented on the following pages, indicate the category of each utterance. Multiple codes may be used.

8. Degree: Indicate how many QRF cycles are contained in the assessment episode. This need only be placed on one line, preferably on the line adjacent to QU1

continued
Appendix H: Coding Book continued

### Codes for Initial Questions

<table>
<thead>
<tr>
<th>R</th>
<th>Rhetorical</th>
</tr>
</thead>
<tbody>
<tr>
<td>K</td>
<td>Knowledge</td>
</tr>
<tr>
<td>C</td>
<td>Comprehension</td>
</tr>
<tr>
<td>A</td>
<td>Application (includes prediction, calculation, interpreting graphs)</td>
</tr>
<tr>
<td>H</td>
<td>Higher-order (analysis, evaluation, synthesis)</td>
</tr>
</tbody>
</table>

### Codes for Student Responses

<table>
<thead>
<tr>
<th>N</th>
<th>none, before teacher intervenes</th>
</tr>
</thead>
<tbody>
<tr>
<td>OT</td>
<td>off-topic response</td>
</tr>
<tr>
<td>NV</td>
<td>non-verbal response</td>
</tr>
<tr>
<td>RR</td>
<td>request for clarification, repetition</td>
</tr>
<tr>
<td>YN</td>
<td>yes/no response</td>
</tr>
<tr>
<td>WP</td>
<td>word or phrase response (includes numerical values)</td>
</tr>
<tr>
<td>E</td>
<td>explanation</td>
</tr>
<tr>
<td>RE</td>
<td>response (WP answer) plus explanation</td>
</tr>
<tr>
<td>T</td>
<td>tangential but on-task response</td>
</tr>
<tr>
<td>IDK</td>
<td>student does not know and says so</td>
</tr>
<tr>
<td>B</td>
<td>responses are displayed on the board or projector</td>
</tr>
<tr>
<td>H</td>
<td>provides a hypothesis or conjecture</td>
</tr>
</tbody>
</table>

### Codes for Subsequent Questions

<table>
<thead>
<tr>
<th>L</th>
<th>leading question/ prompt student to fill in blank</th>
</tr>
</thead>
<tbody>
<tr>
<td>SA</td>
<td>ask to repeat the response</td>
</tr>
<tr>
<td>RS</td>
<td>reflect to class or single student</td>
</tr>
<tr>
<td>RV</td>
<td>repeat verbatim the previous question</td>
</tr>
<tr>
<td>RC</td>
<td>rephrase or clarify the previous question</td>
</tr>
<tr>
<td>RE</td>
<td>request for elaboration (fishing for more information)</td>
</tr>
<tr>
<td>FZ</td>
<td>focus and zoom (tell more about a part of student response) (Collapsed with the RE codes for final analysis)</td>
</tr>
<tr>
<td>I</td>
<td>Implied (uses student name only, provides more wait time)</td>
</tr>
</tbody>
</table>

continued
Appendix H: Coding Book continued

Identifying Evidence of Teacher Awareness of Student Learning From Interviews

1. Read through transcribed interviews. Use NVivo to code for categories. Refer to original audio recordings for clarification of interviewee comments.

2. Any mention related to teacher awareness of student learning, from either the teacher or the student perspective, is included here.

3. Specific questions to look for in interviews include:
   a. Student focus group: Item 5, “How do you think the teacher knows when you understand a concept in this class?”
   b. Post-observation interview: Item 6, “How did you use the TI-Navigator™ specifically to support student learning?”, Item 8, “How might what you learned today about students’ knowledge change what you do in the subsequent lesson?”
   c. Fall telephone interview: None explicitly
   d. Spring telephone interview: Item 10, “Has the use of the TI-Navigator™ supported your understanding of your students’ knowledge of physical science? If so, then how has it supported your understanding?”

4. The entire interview must be analyzed for comments related to teacher awareness of any aspect of student learning, not just the specific items mentioned above.

continued
Appendix H: Coding Book continued

Identifying Changes in Instruction From Interviews

1. Read through transcribed interviews.
2. Include any mention of
3. Specific questions to look for in interviews include:
4. Student focus group: None explicitly
   a. Post-observation interview: Item 8, “How might what you learned today about students’ knowledge change what you do in the subsequent lesson?”; Item 14, “Considering our discussions of the pedagogy for using the TI-Navigator™, on a scale of one to ten how would you rate your implementation of the pedagogy at this point? Can you give an example?”
   b. Fall telephone interview: Item 4, “How did the use of the TI-Navigator™ influence your teaching strategies?”; Item 5, “What examples can you provide of how use of the connected classroom changed your teaching strategies or plans?”
   c. Spring telephone interview: Item 13, “How has using the TI-Navigator™ influenced the way you plan and implement instruction? Can you give an example of changes made during instruction?”
5. The entire interview must be analyzed for comments related to changes in instruction, not just the specific items mentioned above.